

机械工程学院实验报告

Experiment Report

School of Mechanical Engineering

实验课程名称 (Experiment Course)		数值分析和优化算法					
实验项目名称 (Experiment Item)		描述性分析实验					
姓名 (Name)	宋天佑	学号 (No.)	20211414 10279	系别 (Department)	机械设计制造及其自动化	班级 (Class)	2 班
同组学生姓名 (Accompaniers)							
实验日期 (Date)	2025.4.21	实验地点 (place)	CAD 中心 202	实验学时 (Hour)	3		
指导教师姓名 (Superior)	赵秀粉	成绩 (Grade)		指导教师签名 (Superior Signature)			

(以下是实验报告的参考项目，请学生按照实验指导教师的要求在空白处依次填写，不够加附页)

一、实验目的 (Experiment Objectives)

- 1、熟练掌握线性回归模型的建立方法，掌握回归函数及回归分析类的使用方法。
- 2、掌握总离差平方和 SST、回归平方和 SSR、残差平方和 SSE 等相关统计量的意义及编程程序。
- 3、掌握逐步回归的思想与方法，掌握 stepwise 命令的使用方法。
- 4、掌握残差分析方法。

二、实验仪器与设备 (包括名称、型号、规格等) (Main Experiment Apparatus)

安装有 MATLAB 的计算机。

三、实验内容、步骤与结论 (Experiment Contents, Experiment Steps and Result)

实验 1:

```
1 %  
2 % coding: vtf-8 ->  
3 % time : 2025/4/18 19:35  
4 % author : 宋天佑  
5 % file : vtf.m  
6 % software: Visual Studio Code  
7 %  
8 % 输入数据  
9 % 清除  
10 load data1.mat  
11 % = data1(1, :);  
12 y = data1(2, :);  
13 % 输出  
14 Lxx = sum((x - mean(x)).^ 2);  
15 Lxy = sum((x - mean(x)) .* (y - mean(y)));  
16  
17 b1 = Lxy / Lxx;  
18 b0 = mean(y) - mean(x) * b1;  
19  
20 y1 = b0 + b1 * x;  
21 disp(sprintf('y = %f', 1, 100));  
22 disp('问题1: newLine');  
23 % = sprintf('y = %f * x + %f * x1 ', b0, b1);  
24 disp('一元线性回归方程为: ');  
25 disp(s);  
26  
27 % 计算统计量  
28  
29 TSS = sum((y - mean(y)).^ 2);  
30 ESS = sum((y1 - mean(y)).^ 2);  
31 RSS = sum((y - y1).^ 2);  
32 sgnd = RSS / (size(x, 2) - 2);  
33 F = ESS / RSS / (size(x, 2) - 2);  
34 t = Lxy / sqrt(Lxx * sgnd);  
35 R2 = ESS / TSS;  
36  
37 % 显示结果  
38 % = R, R2;  
39 F, reject_domain = fval(1 - a, 1, size(x, 2) - 2); % 显著水平 0.05, 95% 置信区间  
40 L_reject_domain = tval(1 - a/2, size(x, 2) - 2);  
41 t_reject_domain = t_reject_domain(2);  
42  
43 reject_domain = [F, t_R2; F_reject_domain, t_reject_domain, R, R2];  
44 reject_domain_name = {'F 统计量', 't 统计量', '决定系数 R 平方'};
```

```
45 % 显示回归方程  
46 disp(sprintf('y = %f', 1, 100));  
47 disp('问题2: newLine');  
48  
49 for i = 1:size(reject_domain, 2)  
50  
51     if i == 1  
52         % 显示 F 统计量: xx 变量对 y 的线性影响  
53         F1 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
54         s1 = sprintf(F1, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
55         disp(s1);  
56     end  
57  
58     % 显示 t 统计量: xx 变量对 y 的线性影响  
59     for i = 1:size(reject_domain, 2)  
60  
61         if (reject_domain(1, 1) > reject_domain(2, 1)) && i < 3  
62  
63             F2 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
64             s2 = sprintf(F2, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
65             disp(s2);  
66  
67             F3 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
68             s3 = sprintf(F3, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
69             disp(s3);  
70  
71             F4 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
72             s4 = sprintf(F4, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
73             disp(s4);  
74  
75             F5 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
76             s5 = sprintf(F5, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
77             disp(s5);  
78  
79             F6 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
80             s6 = sprintf(F6, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
81             disp(s6);  
82  
83             F7 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
84             s7 = sprintf(F7, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
85             disp(s7);  
86  
87             F8 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
88             s8 = sprintf(F8, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
89             disp(s8);  
90  
91             F9 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
92             s9 = sprintf(F9, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
93             disp(s9);  
94  
95             F10 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
96             s10 = sprintf(F10, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
97             disp(s10);  
98  
99             F11 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
100             s11 = sprintf(F11, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
101             disp(s11);  
102  
103             F12 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
104             s12 = sprintf(F12, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
105             disp(s12);  
106  
107             F13 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
108             s13 = sprintf(F13, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
109             disp(s13);  
110  
111             F14 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
112             s14 = sprintf(F14, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
113             disp(s14);  
114  
115             F15 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
116             s15 = sprintf(F15, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
117             disp(s15);  
118  
119             F16 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
120             s16 = sprintf(F16, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
121             disp(s16);  
122  
123             F17 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
124             s17 = sprintf(F17, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
125             disp(s17);  
126  
127             F18 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
128             s18 = sprintf(F18, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
129             disp(s18);  
130  
131             F19 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
132             s19 = sprintf(F19, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
133             disp(s19);  
134  
135             F20 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
136             s20 = sprintf(F20, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
137             disp(s20);  
138  
139             F21 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
140             s21 = sprintf(F21, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
141             disp(s21);  
142  
143             F22 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
144             s22 = sprintf(F22, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
145             disp(s22);  
146  
147             F23 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
148             s23 = sprintf(F23, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
149             disp(s23);  
150  
151             F24 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
152             s24 = sprintf(F24, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
153             disp(s24);  
154  
155             F25 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
156             s25 = sprintf(F25, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
157             disp(s25);  
158  
159             F26 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
160             s26 = sprintf(F26, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
161             disp(s26);  
162  
163             F27 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
164             s27 = sprintf(F27, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
165             disp(s27);  
166  
167             F28 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
168             s28 = sprintf(F28, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
169             disp(s28);  
170  
171             F29 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
172             s29 = sprintf(F29, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
173             disp(s29);  
174  
175             F30 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
176             s30 = sprintf(F30, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
177             disp(s30);  
178  
179             F31 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
180             s31 = sprintf(F31, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
181             disp(s31);  
182  
183             F32 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
184             s32 = sprintf(F32, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
185             disp(s32);  
186  
187             F33 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
188             s33 = sprintf(F33, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
189             disp(s33);  
190  
191             F34 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
192             s34 = sprintf(F34, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
193             disp(s34);  
194  
195             F35 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
196             s35 = sprintf(F35, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
197             disp(s35);  
198  
199             F36 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
200             s36 = sprintf(F36, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
201             disp(s36);  
202  
203             F37 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
204             s37 = sprintf(F37, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
205             disp(s37);  
206  
207             F38 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
208             s38 = sprintf(F38, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
209             disp(s38);  
210  
211             F39 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
212             s39 = sprintf(F39, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
213             disp(s39);  
214  
215             F40 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
216             s40 = sprintf(F40, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
217             disp(s40);  
218  
219             F41 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
220             s41 = sprintf(F41, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
221             disp(s41);  
222  
223             F42 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
224             s42 = sprintf(F42, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
225             disp(s42);  
226  
227             F43 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
228             s43 = sprintf(F43, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
229             disp(s43);  
230  
231             F44 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
232             s44 = sprintf(F44, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
233             disp(s44);  
234  
235             F45 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
236             s45 = sprintf(F45, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
237             disp(s45);  
238  
239             F46 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
240             s46 = sprintf(F46, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
241             disp(s46);  
242  
243             F47 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
244             s47 = sprintf(F47, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
245             disp(s47);  
246  
247             F48 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
248             s48 = sprintf(F48, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
249             disp(s48);  
250  
251             F49 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
252             s49 = sprintf(F49, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
253             disp(s49);  
254  
255             F50 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
256             s50 = sprintf(F50, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
257             disp(s50);  
258  
259             F51 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
260             s51 = sprintf(F51, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
261             disp(s51);  
262  
263             F52 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
264             s52 = sprintf(F52, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
265             disp(s52);  
266  
267             F53 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
268             s53 = sprintf(F53, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
269             disp(s53);  
270  
271             F54 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
272             s54 = sprintf(F54, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
273             disp(s54);  
274  
275             F55 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
276             s55 = sprintf(F55, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
277             disp(s55);  
278  
279             F56 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
280             s56 = sprintf(F56, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
281             disp(s56);  
282  
283             F57 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
284             s57 = sprintf(F57, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
285             disp(s57);  
286  
287             F58 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
288             s58 = sprintf(F58, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
289             disp(s58);  
290  
291             F59 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
292             s59 = sprintf(F59, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
293             disp(s59);  
294  
295             F60 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
296             s60 = sprintf(F60, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
297             disp(s60);  
298  
299             F61 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
300             s61 = sprintf(F61, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
301             disp(s61);  
302  
303             F62 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
304             s62 = sprintf(F62, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
305             disp(s62);  
306  
307             F63 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
308             s63 = sprintf(F63, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
309             disp(s63);  
310  
311             F64 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
312             s64 = sprintf(F64, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
313             disp(s64);  
314  
315             F65 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
316             s65 = sprintf(F65, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
317             disp(s65);  
318  
319             F66 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
320             s66 = sprintf(F66, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
321             disp(s66);  
322  
323             F67 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
324             s67 = sprintf(F67, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
325             disp(s67);  
326  
327             F68 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
328             s68 = sprintf(F68, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
329             disp(s68);  
330  
331             F69 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
332             s69 = sprintf(F69, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
333             disp(s69);  
334  
335             F70 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
336             s70 = sprintf(F70, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
337             disp(s70);  
338  
339             F71 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
340             s71 = sprintf(F71, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
341             disp(s71);  
342  
343             F72 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
344             s72 = sprintf(F72, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
345             disp(s72);  
346  
347             F73 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
348             s73 = sprintf(F73, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
349             disp(s73);  
350  
351             F74 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
352             s74 = sprintf(F74, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
353             disp(s74);  
354  
355             F75 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
356             s75 = sprintf(F75, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
357             disp(s75);  
358  
359             F76 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
360             s76 = sprintf(F76, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
361             disp(s76);  
362  
363             F77 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
364             s77 = sprintf(F77, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
365             disp(s77);  
366  
367             F78 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
368             s78 = sprintf(F78, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
369             disp(s78);  
370  
371             F79 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
372             s79 = sprintf(F79, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
373             disp(s79);  
374  
375             F80 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
376             s80 = sprintf(F80, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
377             disp(s80);  
378  
379             F81 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
380             s81 = sprintf(F81, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
381             disp(s81);  
382  
383             F82 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
384             s82 = sprintf(F82, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
385             disp(s82);  
386  
387             F83 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
388             s83 = sprintf(F83, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
389             disp(s83);  
390  
391             F84 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
392             s84 = sprintf(F84, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
393             disp(s84);  
394  
395             F85 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
396             s85 = sprintf(F85, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
397             disp(s85);  
398  
399             F86 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
400             s86 = sprintf(F86, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
401             disp(s86);  
402  
403             F87 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
404             s87 = sprintf(F87, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
405             disp(s87);  
406  
407             F88 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
408             s88 = sprintf(F88, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
409             disp(s88);  
410  
411             F89 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
412             s89 = sprintf(F89, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
413             disp(s89);  
414  
415             F90 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
416             s90 = sprintf(F90, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
417             disp(s90);  
418  
419             F91 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
420             s91 = sprintf(F91, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
421             disp(s91);  
422  
423             F92 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
424             s92 = sprintf(F92, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
425             disp(s92);  
426  
427             F93 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
428             s93 = sprintf(F93, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
429             disp(s93);  
430  
431             F94 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
432             s94 = sprintf(F94, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
433             disp(s94);  
434  
435             F95 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
436             s95 = sprintf(F95, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
437             disp(s95);  
438  
439             F96 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
440             s96 = sprintf(F96, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
441             disp(s96);  
442  
443             F97 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
444             s97 = sprintf(F97, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
445             disp(s97);  
446  
447             F98 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
448             s98 = sprintf(F98, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
449             disp(s98);  
450  
451             F99 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
452             s99 = sprintf(F99, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
453             disp(s99);  
454  
455             F100 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
456             s100 = sprintf(F100, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
457             disp(s100);  
458  
459             F101 = 'F 统计量为 %f, 95% 置信区间为 %f 到 %f';  
460             s101 = sprintf(F101, reject_domain_name(1), reject_domain(1, 1), (1 - a)*100, reject_domain_name(1), reject_domain(1, 1));  
461             disp(s
```

## 主函数\_1

```
90 end
91
92 y_predict = b0 + b1 * 180;
93
94 disp(repmat('#', 1, 100))
95 disp(['问题3' newline])
96 disp(['若在XX市购买180平方米的房屋，估计购房费用为 ' num2str(y_predict)])
97
98 disp(repmat('#', 1, 100))
```

## 主函数\_2

```
#####
问题1
一维线性回归方程为：
y = -22.285714 + 0.753571 * x1
#####
问题2
F检验值为 2.534787，95%置信度的F检验临界值为 5.987378
t检验值为 9.552609，95%置信度的t检验临界值为 2.446912
在显著性水平0.05下，本次回归未通过F检验，表明自变量对因变量的线性关系不显著
在显著性水平0.05下，本次回归通过t检验，表明自变量对因变量的线性关系显著
可决系数R2检验的值：0.938305 > 0.800000，故本次回归的拟合程度较好
#####
问题3
若在XX市购买180平方米的房屋，估计购房费用为 113.3571
#####
```

## 主函数\_3

## 输出

## 实验 2:

```
1 %{\n 2 *- coding: utf-8 -*\n 3 @time : 2025/4/18 19:35\n 4 @Author : 熊Sol\n 5 @File : e3.m\n 6 @Software: Visual Studio Code\n 7 %}\n 8 clc, clear\n 9\n10 %\n多项式回归\n11 load data2.mat\n12 x = data2(1, :);\n13 y = data2(2, :);\n14\n15 % 有bug，这里拟合使用和[p,s]算出的p都是正确的，但[p,s,mu]的值是错误的\n16 [p, s, mu] = polyfit(x, y, 2);\n17\n18 x_test = linspace(0, 20, 100);\n19 y_predict = polyval(p, x_test);\n20 % 画图(可以不用画)\n21 s = sprintf('y = %.3fx^2 + %.3fx + %.3f', p(1), p(2), p(3));\n22\n23 % set(gcf, 'unit', 'normalized', 'position', [0.2, 0.2, 0.5, 0.5])\n24\n25 % plot(x, y, 'bo', x_test, y_predict, 'r-')\n26\n27 % grid on\n28\n29\n30\n31 % str = {'拟合曲线y关于横坐标x的二次多项式回归方程:'};\n32\n33 % text(2, 0.0, str, 'FontSize', 20, 'FontName', '宋体', 'Color', 'black')\n34\n35 % legend({'原数据', '二次多项式回归拟合图'}, 'FontName', '宋体', 'FontSize', 15, 'Location', 'west', 'Position', [0.2 0.5 0.1\n36 disp('多项式回归的方程为: ')\n37 disp(s)
```

多项式回归的方程为：  
 $y = 0.140x^2 + 0.197x + 1.010$

## 主函数

## 输出

## 实验 3:

```
1 %{\n 2 *- coding: utf-8 -*\n 3 @time : 2025/4/18 19:35\n 4 @Author : 熊Sol\n 5 @File : e3.m\n 6 @Software: Visual Studio Code\n 7 %}\n 8 clc, clear\n 9\n10 %\n导入数据\n11 load data3.mat\n12\n13 value_name = {'序号', '反应速度 y', '氢 x1', 'n戊烷 x2', '异构戊烷 x3'};\n14 data3_table = array2table(data3, "VariableNames", value_name);\n15\n16 % 回归计算中，可以使用表格数据计算，但数值矩阵比较方便，和平常的操作一样\n17\n18 X = data3_table(:, 3:end); % 使用{}提取数值数据\n19 y = data3_table(:, 2);\n20\n21 %\n建立模型\n22 model_func = @(b, X) (b(1) .* X(:, 2) - X(:, 3) / b(5)) ./ ...\n23 (1 + b(2) .* X(:, 1) + b(3) .* X(:, 2) + b(4) .* X(:, 3));\n24\n25 % 初始参数\n26 beta0 = ones(1, 5);\n27\n28 %\n非线性回归\n29 mdl = fitnlm(X, y, model_func, beta0);\n30\n31 %\n显示结果\n32 result = mdl.coefCI;\n33 disp(repmat('#', 1, 100))\n34 disp('b1-b5的置信区间')\n35 % b1的置信区间为 xxx - xxx\n36 for i = 1:5\n37     formatSpec = "b%d的置信区间为 %.4f --- %.4f";\n38     str = sprintf(formatSpec, i, result(i, 1), result(i, 2));\n39     disp(str)\n40 end\n41\n42 disp(repmat('#', 1, 100))
```

```
#####\nb1-b5的置信区间\nb1的置信区间为 -0.7468 --- 3.2520\nb2的置信区间为 -0.0377 --- 0.1632\nb3的置信区间为 -0.0312 --- 0.1113\nb4的置信区间为 -0.0609 --- 0.2857\nb5的置信区间为 -0.7381 --- 3.1208\n#####
```

## 主函数

## 输出

## 实验 4:

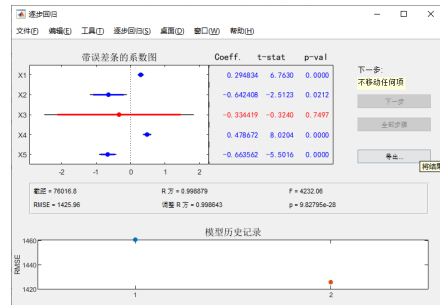
```
1 %f
2 % coding: utf-8 -*-
3 %time : 2025/4/18 19:35
4 %author : 周立
5 %file : ex4.m
6 %Software: Visual Studio Code
7 %
8 %c; clear
9 % 导入数据
10 load data4.mat
11 value_name = {'x1', 'x2', 'x3', 'x4', 'x5', 'y'};
12 data4_table = array2table(data4, 'VariableNames', value_name);
13
14 X = data4_table(:, 2:end - 1);
15 y = data4_table(:, end);
16 % 1.1 多元线性回归 不含交互项
17
18 [r, c] = size(X);
19 mdl = fitlm(X, y, 'y=x1+x2+x3+x4+x5');
20
21 % 1.2 % 输出参数
22 % 回归系数
23 beta = mdl.Coefficients(:, 1);
24
25 disp(repmat(' ', 1, 100));
26 disp('多元线性回归回归方程为:');
27 s = sprintf('y = %f + %f * x1 + %f * x2 + %f * x3 + %f * x4 + %f * x5', beta(1), beta(2), beta(3), beta(4), beta(5));
28 disp(s);
29 % 1.3 显著性检验
30 t_pvalue = mdl.Coefficients(:, 'pvalue');
31 t_pvalue = t_pvalue(2:end);
32
33 summary = anova(mdl, 'summary');
34 f_pvalue = summary(2, 'pvalue');
35 disp(repmat(' ', 1, 100));
36 % F 检验
37 if t_pvalue < 0.05
38     disp('模型整体具有统计显著性(a=0.05水平下), 即至少有一个自变量与因变量之间存在显著的线性关系。');
39 else
40     disp('模型整体不具有统计显著性(显著性水平a=0.05水平下), 即自变量与因变量之间不存在显著的线性关系。');
41 end
42
```

```
43 % t 检验
44 t_logical = t_pvalue < 0.5;
45 t_logical_true = [false; t_logical];
46 t_logical_false = [false; ~t_logical];
47
48 disp(repmat(' ', 1, 100));
49 disp('在显著性水平a=0.05下, 以下自变量对因变量的影响具有统计显著性:');
50 s = strjoin(value_name(t_logical_true));
51 disp(s);
52 disp('在显著性水平a=0.05下, 以下自变量对因变量的影响不具有统计显著性:');
53 s = strjoin(value_name(t_logical_false));
54 disp(s);
55 % R2
56 disp(repmat(' ', 1, 100));
57 adjust_R2 = mdl.Rsquared.Adjusted;
58 R2 = mdl.Rsquared.Ordinary;
59 s = sprintf('R2的值为 %f, 调整R2的值为 %f, 具有非常强的线性关系。', R2, adjust_R2);
60 disp(s);
61
62 disp(repmat(' ', 1, 100));
63
64 % 2 逐步回归
65 stepwise(X, y, 1:5, 0.05, 0.10);
66
```

### 主函数\_1

```
#####
多元线性的回归方程为:
y = 82742.407595 + 0.315405 * x1 + -0.594418 * x2 + -0.334419 * x3 + 0.511921 * x4 + -0.724369 * x5
#####
模型整体具有统计显著性(a=0.05水平下), 即至少有一个自变量与因变量之间存在显著的线性关系。
#####
在显著性水平a=0.05下, 以下自变量对因变量的影响具有统计显著性:
x1 x2 x4 x5
在显著性水平a=0.05下, 以下自变量对因变量的影响不具有统计显著性:
x3
#####
R2的值为 0.998885, 调整R2的值为 0.998576, 具有非常强的线性关系。
#####
```

### 主函数\_2



### 输出\_1

### 输出\_2

## 四、实验思考题

### 1、如何判断拟合效果是否满足要求？

通过可决系数  $R^2$  和调整的可决系数 ( $Adjusted R^2$ ) 评估解释能力。  $R^2 = 1 - \frac{SST}{SSR}$  ,  $Adjusted R^2 = 1 - (1 - R^2) \frac{n-1}{n-p-1}$  。均方根误差 ( $RMSE = \sqrt{\frac{SSR}{n}}$ ) 衡量预测精度。残差分析 (残差图、正态概率图) 检验模型假设。统计检验 (F 检验整体显著性, t 检验系数显著性) 评估统计意义。

### 2、在何种情况下进行非线性回归分析？

当因变量与自变量关系无法用线性函数有效描述时, 例如: 理论模型非线性, 散点图呈现曲线, 线性回归模型诊断不通过, 变量变换无法线性化, 或需要拟合特定非线性函数形式。

### 3、逐步回归的思想是什么? 如何实现逐步回归?

思想: 迭代地引入或剔除自变量, 根据统计准则构建最优模型, 平衡复杂性和拟合优度。

实现：

向前选择： 从空模型开始，逐步加入最显著的自变量；

向后剔除： 从全模型开始，逐步剔除最不显著的自变量；

双向逐步回归： 结合向前选择和向后剔除，每步选择后都检验已入选变量的显著性；

#### 4、采用逐步回归进行实际问题解决后，进行回归诊断的标准是什么？

与一般多元线性回归诊断标准相同：**线性性、独立性**（Durbin-Watson 统计量）、**同方差性、正态性、多重共线性（VIF）、异常值和强影响点**（学生化残差、Cook 距离、帽子矩阵）、**模型整体和参数显著性**。

#### 5、残差分析法包括哪些方面？

残差的定义和计算： $e_i = y_i - \hat{y}_i$ ；

残差图分析： 残差 vs. 预测值、自变量、时间序列图，检查随机性、非线性、异方差、自相关；

残差的统计性质分析： 均值、方差、分布（Q-Q 图、统计检验）；

识别异常值和强影响点： 标准化/学生化残差、Cook 距离、帽子矩阵；

检验残差的自相关性： Durbin-Watson 统计量、ACF 和 PACF 图；

### 五、实验收获与建议

本次实验收获颇丰，不仅深化了回归分析理论，更熟练运用了 MATLAB。手动计算统计量（R<sup>2</sup>, F, T）让我从数值层面理解模型优劣和变量显著性。掌握多项式和多元线性回归，体会到模型选择的重要性。自定义非线性回归拓展了处理复杂关系的能力。学习逐步回归，理解了多变量模型构建的策略和权衡。

MATLAB 的强大功能和便捷性令人印象深刻，提升了编程和独立分析能力。实验过程也体会到严谨细致的重要性。

对实验的建议：增加案例背景介绍，更深入探讨模型诊断和改进方法，鼓励更深入的探索思考，加强铜须门之间的交流与合作。

本次实验是一次宝贵的学习经历，提升了理论和实践技能，并引发了更深层次的思考。希望未来实验项目更加完善。