数学实验：第十次作业

计算机系　计 73　2017011620　李家昊

2020 年 5 月 19 日

1　实验目的

- 了解回归分析的基本原理，掌握 MATLAB 实现的方法。
- 练习用回归分析解决实际问题。

2　问题求解

2.1　Chap13-Ex7 有氧锻炼（计算题）

2.1.1　算法设计

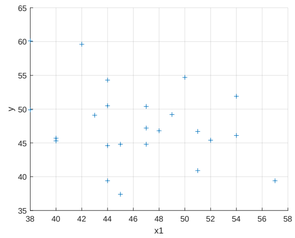
第 (1) 问：一元回归　对于一元回归，首先应当分别画出 y 与 x_1, x_2, x_3, x_4, x_5 的散点图，初步观察它们的相关性，如图 1。可以看出，y 关于 x_3 的图像呈现出明显的负相关，关于 x_4 和 x_5 呈现出较弱的负相关。进一步，可对每个自变量分别进行 n 次多项式回归，即，

$$y = \sum_{i=0}^{n} \beta_i x^i + \varepsilon \tag{1}$$

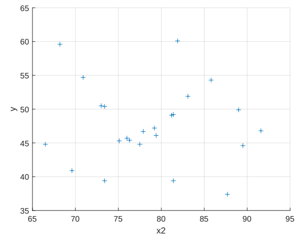
其中，β_i 为回归系数，x 可取 x_1, x_2, x_3, x_4, x_5 中的任意一个。注意到 5 个散点图均不具有明显的高次关系，因此线性回归模型应当足够表达因变量与自变量的相关性，可简化建模为，

$$y = \beta_0 + \beta_1 x + \varepsilon \tag{2}$$

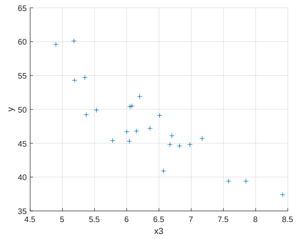
在 MATLAB 中，可通过 `regress` 命令进行线性回归，通过 `polytool` 命令进行多项式回归。



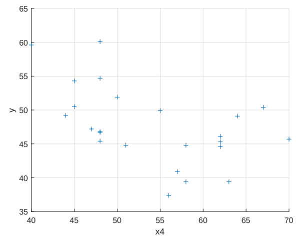
(a) y-x_1 图像



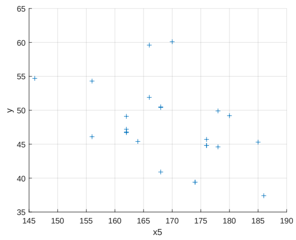
(b) y-x_2 图像



(c) y-x_3 图像



(d) y-x_4 图像



(e) y-x_5 图像

图 1: 因变量 y 关于自变量 x_1, x_2, x_3, x_4, x_5 的散点图像

第 (2) 问：二元回归 对于二元回归，考虑到高次项的贡献较低，而交叉项不具有明确的实际意义，因此只进行多元线性回归，

$$y = \beta_0 + \beta_1 x_i + \beta_2 x_j + \varepsilon \tag{3}$$

其中 $i, j \in \{1, 2, 3, 4, 5\}$ 且规定 $i < j$，遍历 i, j 所有的组合，共 $C_5^2 = 10$ 组，对每组 i, j 进行多元线性回归，选出剩余方差 s^2 最小的一组，作为最优的二元回归模型。

在 MATLAB 中，可使用 `stepwise` 命令进行逐步回归，使用 `regress` 命令进行最终计算。

第 (3) 问：多元回归 对于多元回归，仍然采用多元线性回归的方式，

$$y = \beta_0 + \sum_{i=1}^{5} \beta_i x_i + \varepsilon \tag{4}$$

在 MATLAB 中，同样采用 `stepwise` 进行逐步回归的探索，使用 `regress` 命令进行多元线性回归的计算。

第 (4) 问：去除异常点 残差的置信区间不含零点，则为异常点。在 MATLAB 中，可通过 `rcoplot` 画出残差图，异常点将被标记为红色，然后逐步剔除异常点即可。

2.1.2 程序

请参见附录4.1。

2.1.3 计算结果

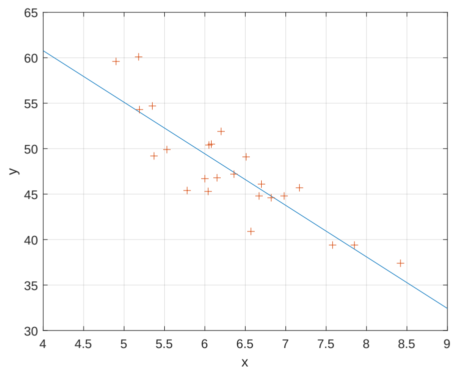
第 (1) 问：一元回归 经过回归分析，计算结果如表 1，可以看出，在默认的显著性水平 $\alpha = 0.05$ 下，选择 x_3, x_4, x_5 时均满足 $p < \alpha$，可得到有效的回归模型，其中，当选择 x_3 时，决定系数 R^2 最大，p 值最小，因此应当选择 x_3 作为自变量，

$$y = 83.44 - 5.67x_3 + \varepsilon \tag{5}$$

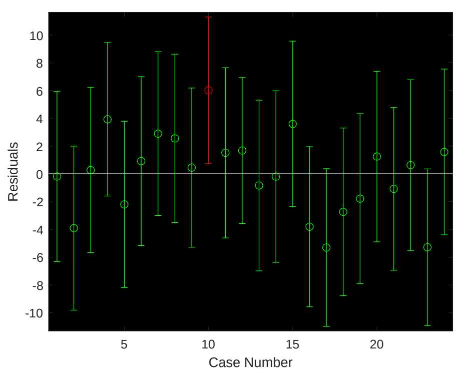
方程 (5) 即为最优的一元线性回归模型，此时画出拟合图像和残差图像如图 2。

表 1: 一元线性回归模型

自变量	$\hat{\beta}_0$	$\hat{\beta}_1$	β_1 置信区间	R^2	F	p	s^2
x_1	64.38	-0.36	[-0.83, 0.11]	0.1025	2.5115	0.1273	31.2484
x_2	52.80	-0.07	[-0.43, 0.30]	0.0060	0.1337	0.7181	34.6053
x_3	83.44	-5.67	[-7.13, -4.21]	0.7474	65.0959	0.0000	8.7943
x_4	67.11	-0.36	[-0.63, -0.09]	0.2631	7.8560	0.0104	25.6547
x_5	94.00	-0.27	[-0.51, -0.04]	0.2091	5.8169	0.0247	27.5352



(a) 线性拟合



(b) 残差图像

图 2: 一元最优线性回归模型

为了进一步探索高次回归模型，尝试对自变量 x_3 进行 2 次和 3 次多项式回归，得到各回归系数的置信区间如表 2。可以看出，当采用 2 次或 3 次多项式回

归时，最高项系数的置信区间均包含零点，不是有效的回归模型。因此，线性回归模型方程 (5) 即为最优的一元回归模型。

表 2: 一元多项式回归模型

回归模型	β_0 置信区间	β_1 置信区间	β_2 置信区间	β_3 置信区间
一次多项式	[74.16, 92.72]	[-7.13, -4.21]	/	/
二次多项式	[67.19, 178.26]	[-35.04, -0.78]	[-0.37, 2.24]	/
三次多项式	[11.62, 803.26]	[-334.88, 33.02]	[-6.79, 49.47]	[-2.44, 0.39]

第 (2) 问：二元回归 经过逐步回归和最终计算，得到剩余方差 s^2 最小的二元线性回归模型为，

$$y = 90.85 - 0.19x_1 - 5.47x_3 + \varepsilon \tag{6}$$

方程 (6) 即为最优的二元回归模型。详细参数如表 3，可以看出，回归系数 β_1 的置信区间包含零点，表明 x_1 的引入对模型的拟合效果贡献不大。

表 3: 二元线性回归模型

回归系数	估计值	置信区间
β_0	90.8529	[77.5587, 104.1471]
β_1	-0.1870	[-0.4337, 0.0598]
β_2	-5.4671	[-6.9059, -4.0283]
$R^2 = 0.7741,\quad F = 35.9833,\quad p = 0.0000,\quad s^2 = 8.2389$		

第 (3) 问：多元回归 经过逐步回归和最终计算，得到剩余方差 s^2 最小的多元线性回归模型为，

$$y = 118.01 - 0.33x_1 - 4.57x_3 - 0.16x_5 + \varepsilon \tag{7}$$

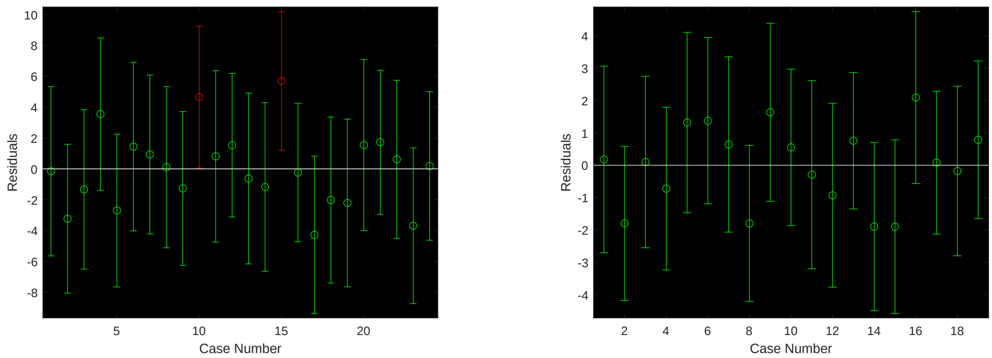
方程 (7) 即为最优的多元回归模型。详细参数如表 4。

表 4: 多元线性回归模型

回归系数	估计值	置信区间
β_0	118.0135	[88.1010, 147.9260]
β_1	-0.3254	[-0.5940, -0.0568]
β_3	-4.5694	[-6.1842, -2.9546]
β_5	-0.1561	[-0.3126, 0.0004]
$R^2 = 0.8143,\quad F = 29.2364,\quad p = 0.0000,\quad s^2 = 7.1112$		

在一元回归、二元回归和多元回归的模型中，多元回归模型方程 (7) 的剩余方差 s^2 最小，决定系数 R^2 最大，因此将其作为最终模型。

第 (4) 问：去除异常点 对最终模型方程 (7) 进行回归时，画出残差图如图 3(a)，其中有两个异常点，去除后又产生了新的异常点，逐步去除所有异常点，序号为 4,10,15,17,23，得到的残差图如图 3(b)。



(a) 去除前 (b) 去除后

图 3: 去除异常点前后的残差图

去除异常点后，模型参数也有了相应的改变，修正后的模型为，

$$y = 109.35 - 0.22x_1 - 3.77x_3 - 0.17x_5 + \varepsilon \tag{8}$$

详细参数如表 5。可以看出，去除异常点后，决定系数 R^2 和剩余方差 s^2 都有了明显的改善。

表 5: 去除异常点后的最终模型

回归系数	估计值	置信区间
β_0	109.3470	[92.5520, 126.1419]
β_1	-0.2230	[-0.3912, -0.0548]
β_3	-3.7694	[-4.7072, -2.8316]
β_5	-0.1652	[-0.2483, -0.0821]
$R^2 = 0.9268$, $F = 63.3391$, $p = 0.0000$, $s^2 = 1.8553$		

2.1.4 结果分析

从最终模型方程 (8) 可以看出，耗氧能力 y 与年龄 x_1、1500 米跑用时 x_3、以及跑步后心速 x_5 均呈负相关，即年龄越大、1500 米跑得越慢、跑步后心速越快的人，耗氧能力越低，该结论与生活经验相符。

2.1.5 结论

若 $x_1 \sim x_5$ 中只许选择 1 个变量，最好的模型是方程 (5)。

若 $x_1 \sim x_5$ 中只许选择 2 个变量，最好的模型是方程 (6)。

若不限制变量个数，最好的模型是方程 (7)。选择方程 (7) 作为最终模型，因为该模型的剩余方差最小。

最终模型的残差图如图 3(a)，共有 5 个异常点，剔除所有异常点后，模型修正为方程 (8)。

2.2 Chap13-Ex9 洗衣粉泡沫（计算题）

2.2.1 算法设计

第 (1) 问 将搅拌程度 x_1 作为普通变量，建立 y 与 x_1 和 x_2 的多元线性回归模型为，

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon \tag{9}$$

在 MATLAB 中，可使用 `stepwise` 命令进行逐步回归，使用 `regress` 命令计算多元线性回归。

第 (2) 问 为了避免引入定量关系，可用两个变量 (x_{1a}, x_{1b}) 表达搅拌程度 x_1，令 $(0,0)$ 表示 $x_1 = 1$，令 $(0,1)$ 表示 $x_1 = 2$，令 $(1,0)$ 表示 $x_1 = 3$，可建立 y 与 x_{1a}, x_{1b}, x_2 的多元线性回归模型为，

$$y = \beta_0 + \beta_{1a} x_{1a} + \beta_{1b} x_{1b} + \beta_2 x_2 + \varepsilon \tag{10}$$

同样使用 `stepwise` 和 `regress` 命令求解。

第 (3) 问 在第 (2) 问的基础上，加入搅拌程度与洗衣粉用量的交互项，可建模如下，

$$y = \beta_0 + \beta_{1a} x_{1a} + \beta_{1b} x_{1b} + \beta_2 x_2 + \beta_{1a2} x_{1a} x_2 + \beta_{1b2} x_{1b} x_2 + \varepsilon \tag{11}$$

同样使用 `stepwise` 和 `regress` 命令求解。

2.2.2 程序

请参见附录4.2。

2.2.3 计算结果

第 (1) 问 经过逐步回归和最终计算，得到剩余方差 s^2 最小的多元线性回归模型为，

$$y = -12.74 + 26.30 x_1 + 3.09 x_2 + \varepsilon \tag{12}$$

详细参数如表 6，得到的残差图像如图 4，从残差图可以看出，当搅拌程度 $x_1 = 2$ 时，残差均为正，其余情况下，残差均为负。表明将搅拌程度的三个水平定量表达时，泡沫高度 y 与搅拌程度 x_1 的线性程度较差。

表 6: 多元线性回归模型

回归系数	估计值	置信区间
β_0	-12.7400	[-29.0268, 3.5468]
β_1	26.3000	[23.1059, 29.4941]
β_2	3.0867	[1.2426, 4.9308]
$R^2 = 0.9654$, $F = 167.5754$, $p = 0.0000$, $s^2 = 21.4910$		

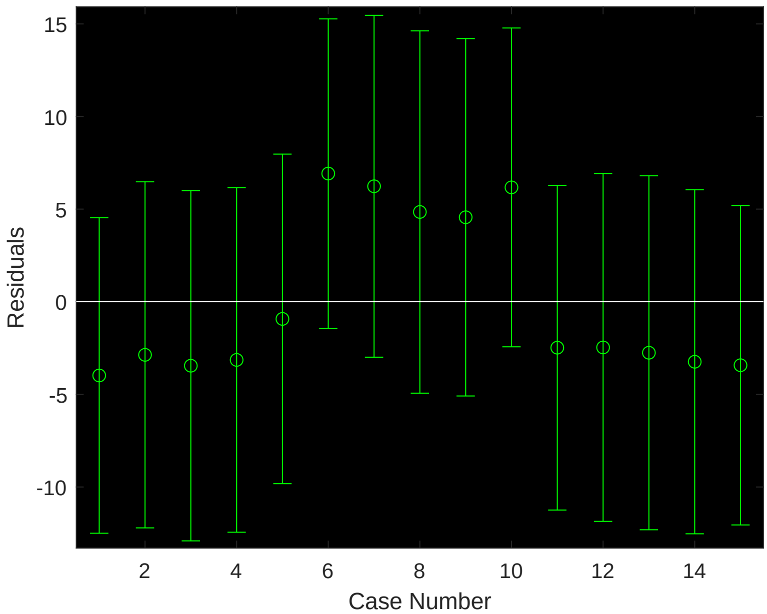


图 4: 残差图像

第 (2) 问 经过逐步回归和最终计算，得到剩余方差 s^2 最小的多元线性回归模型为，

$$y = 10.69 + 52.60x_{1a} + 34.92x_{1b} + 3.09x_2 + \varepsilon \tag{13}$$

详细参数如表 7，与第 (1) 问相比，剩余方差 s^2 显著下降，决定系数 R^2 也有了明显的提升，显然，该模型的拟合效果比第 (1) 问更好，原因是避免了不恰当的定量关系，用 0-1 变量处理不同等级的搅拌程度。

表 7: 多元线性回归模型

回归系数	估计值	置信区间
β_0	10.6867	[7.4475, 13.9259]
β_{1a}	52.6000	[51.2588, 53.9412]
β_{1b}	34.9200	[33.5788, 36.2612]
β_2	3.0867	[2.6995, 3.4738]
$R^2 = 0.9986$, $F = 2675.4529$, $p = 0.0000$, $s^2 = 0.9282$		

得到的残差图如图 5，从残差图可以看出，第 5 个数据是一个异常点，需要进行剔除。

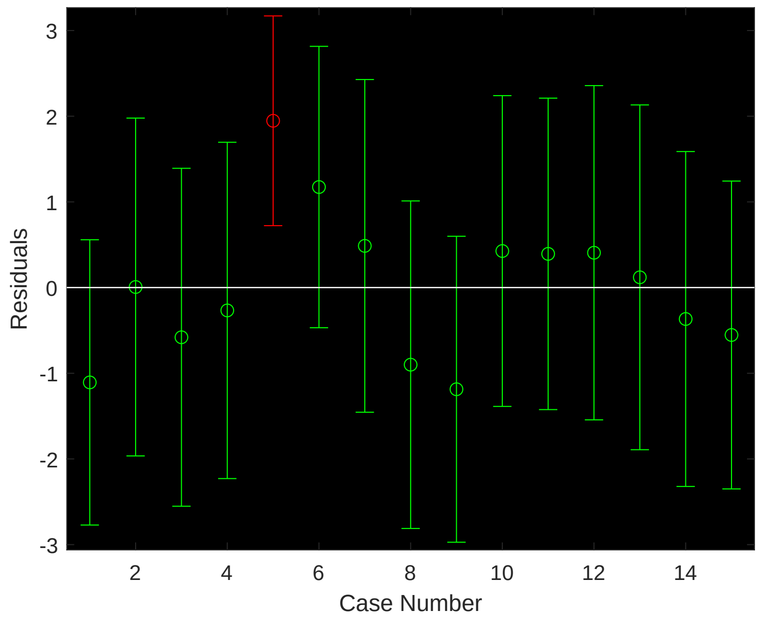


图 5: 残差图像

剔除所有异常点后，模型修正为，

$$y = 11.66 + 53.18x_{1a} + 35.50x_{1b} + 2.89x_2 + \varepsilon \tag{14}$$

此时残差图如图 6。

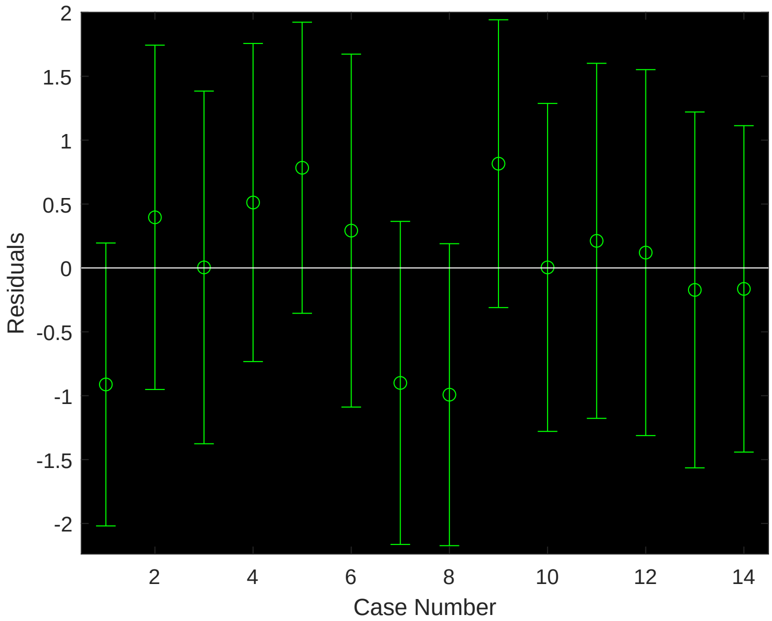


图 6: 残差图像

第 (3) 问 经过计算，得到剩余方差 s^2 最小的多元线性回归模型为，

$$y = 6.02 + 59.40x_{1a} + 42.12x_{1b} + 3.67x_2 - 0.85x_{1a}x_2 - 0.90x_{1b}x_2 + \varepsilon \tag{15}$$

详细参数如表 8，可见，引入交互项后，剩余方差 s^2 和决定系数 R^2 有了进一步改善，模型的拟合效果更好。

表 8: 带有交互项的回归模型

回归系数	估计值	置信区间
β_0	6.0200	[1.6478, 10.3922]
β_{1a}	59.4000	[53.2167, 65.5833]
β_{1b}	42.1200	[35.9367, 48.3033]
β_2	3.6700	[3.1318, 4.2082]
β_{1a2}	-0.8500	[-1.6111, -0.0889]
β_{1b2}	-0.9000	[-1.6611, -0.1389]
$R^2 = 0.9993$, $F = 2634.4605$, $p = 0.0000$, $s^2 = 0.5660$		

2.2.4 结果分析

从方程 (12) 可以看出，x_1, x_2 的回归系数均为正数，表明泡沫高度 y 与搅拌程度 x_1 和洗衣粉用量 x_2 均呈正相关，这与生活经验相符。

从方程 (13) 可以看出，弱搅拌程度对泡沫高度的贡献为 0，中等搅拌程度的贡献为 34.92，强搅拌程度的贡献为 52.60，并不是严格的线性关系，因此，将

三个等级的搅拌程度定量为 $x_1 = 1, 2, 3$ 并不能表达出这种关系，而应当定量为 $x_1 = 0, 34.92, 52.60$，这样就能得到与 0-1 变量相同的结果。然而，这是求解之后才得到的定量方法，在求解之前，使用 0-1 变量来处理会更合适。

2.2.5 结论

将搅拌程度作为普通变量时，回归模型为方程 (12)，残差图上，中等搅拌程度的残差与其他情况明显不同。

将搅拌程度视为没有定量关系的 3 个水平，用 0-1 变量表示时，回归模型为方程 (13)，拟合效果得到了明显改善，但残差图上存在异常点。

加入搅拌程度与洗衣粉用量的交互项，回归模型为方程 (15)，效果更好。

2.3 Chap13-Ex13 Logistic 模型（计算题）

2.3.1 算法设计

第 (1) 问 Logistic 增长曲线模型如下，

$$y_t = \frac{L}{1 + ae^{-kt}} \tag{16}$$

当参数 L, a, k 均未知时，该模型不是一个可线性化模型。当给定 $L = 3000$ 时，则是一个可线性化模型，可转化为，

$$\ln\left(\frac{L}{y_t} - 1\right) = \ln a - kt \tag{17}$$

其中 a, k 为待估参数。上式左端为因变量，记为 y，自变量为 t，令 $\beta_0 = \ln a$ 且 $\beta_1 = -k$，则上式可表示为 $y = \beta_0 + \beta_1 t$。在 MATLAB 中，可采用 `regress` 命令进行线性回归。

第 (2) 问 利用第 (1) 问所得到的 a 和 k 的估计值和 $L = 3000$ 作为 Logistic 模型的拟合初值，对 Logistic 模型做非线性回归。在 MATLAB 中，可采用 `nlinfit` 命令进行非线性回归。

第 (3) 问 Gompertz 模型定义如下，

$$y_t = Le^{-be^{-kt}} \tag{18}$$

取初值 $L_0 = 3000, b_0 = 30, k_0 = 0.4$，用 `nlinfit` 命令进行非线性回归分析。

2.3.2 程序

请参见附录4.3。

2.3.3 计算结果

第 (1) 问 经过计算，得到回归模型参数如表 9。

表 9: Logistic 回归模型

回归系数	估计值	置信区间
β_0	3.8032	[3.5765, 4.0299]
β_1	-0.4941	[-0.5262, -0.4621]
$R^2 = 0.9905$, $F = 1150.7545$, $p = 0.0000$, $s^2 = 0.0386$		

计算参数 a, k 的估计值，

$$\hat{a} = e^{\hat{\beta}_0} = 44.8463, \quad \hat{k} = -\hat{\beta}_1 = 0.4941 \tag{19}$$

画出拟合曲线，如图 7，经过计算，其剩余标准差为 $s = 5053.0508$。

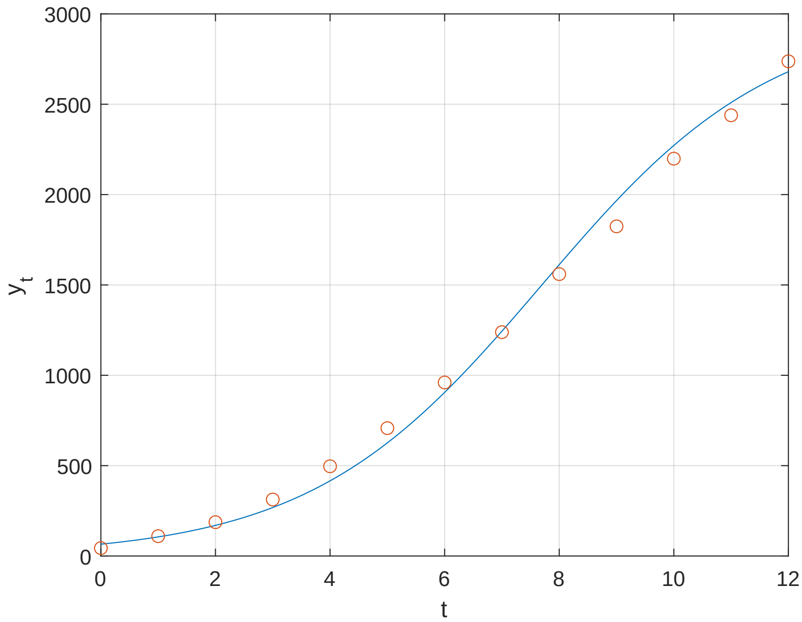


图 7: Logistic 线性回归模型拟合曲线

第 (2) 问 由第 (1) 问计算结果，可得非线性拟合的初值为 $L_0 = 3000, a_0 = 44.8463, k_0 = 0.4941$，经过非线性回归，得到回归系数为，

$$\hat{L} = 3260.4185, \quad \hat{a} = 30.5351, \quad \hat{k} = 0.4148 \tag{20}$$

画出拟合曲线，如图 8，其剩余标准差为 $s = 1765.1289$，模型的拟合效果比线性回归模型更好。

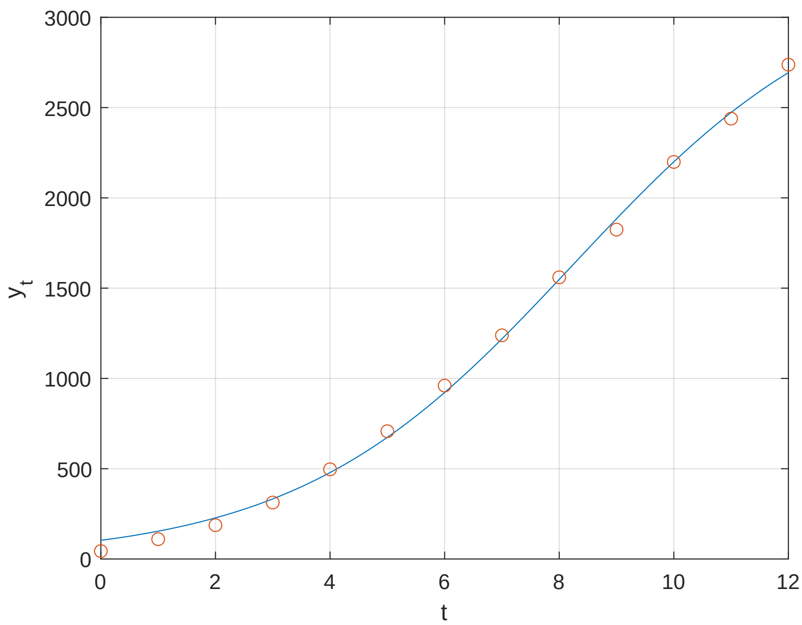


图 8: Logistic 非线性回归模型拟合曲线

第 (3) 问 以 $L_0 = 3000, b_0 = 30, k_0 = 0.4$ 作为初值，经过非线性回归，得到 Gompertz 模型参数的估计值为，

$$\hat{L} = 4810.1269, \quad \hat{b} = 4.5920, \quad \hat{k} = 0.1747 \tag{21}$$

画出拟合曲线，如图 9，其剩余标准差为 $s = 308.1378$，低于 Logistic 模型，说明 Gompertz 模型比 Logistic 模型的拟合效果更好。

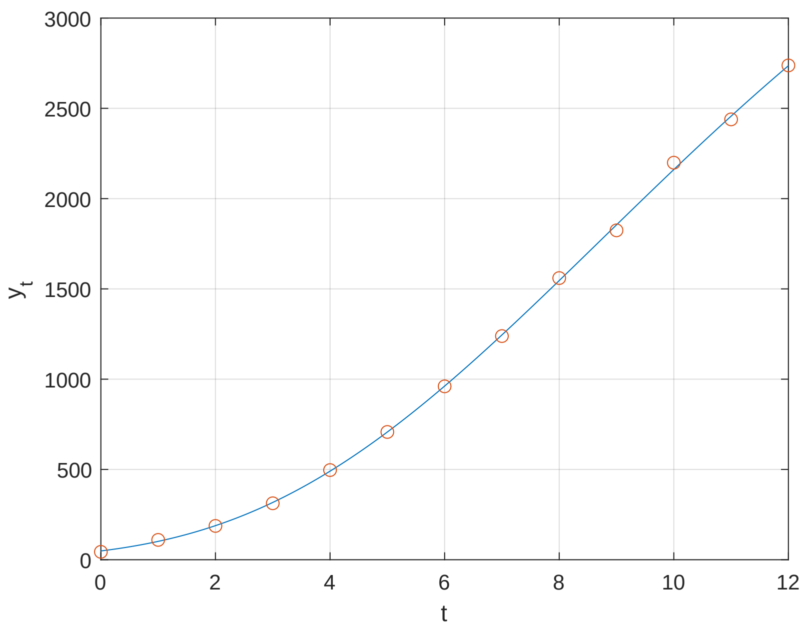


图 9: Gompertz 非线性回归模型拟合曲线

2.3.4 结果分析

求解非线性回归模型时，选取合适的初值，可以降低计算量，加快收敛速度。如果模型可线性化，可以先求出线性化回归模型的参数估计值，将其作为初值，进一步求解非线性回归模型。

2.3.5 结论

Logistic 增长曲线模型不是一个可线性化模型，如果给定 $L = 3000$，则是一个可线性化模型，用线性化模型给出参数 a 和 k 的估计值为 $\hat{a} = 44.8463$ 和 $\hat{k} = 0.4941$。

将上述 a 和 k 的估计值和 $L = 3000$ 作为 Logistic 模型的拟合初值，对 Logistic 模型做非线性回归，得到 $\hat{L} = 3260.4185, \hat{a} = 30.5351, \hat{k} = 0.4148$。

取初值 $L_0 = 3000, b_0 = 30, k_0 = 0.4$，拟合 Gompertz 模型，得到 $\hat{L} = 4810.1269, \hat{b} = 4.5920, \hat{k} = 0.1747$，与 Logistic 模型相比，Gompertz 模型的拟合效果更好。

3 收获与建议

在本次实验中，我掌握了回归分析的基本原理，用回归分析方法建立了实际问题的模型，并用 MATLAB 进行求解，在解决实际问题的过程中，我对数学方法的原理和应用有了更深刻的理解。

感谢老师和助教在这个学期的付出，我在这门课上收获颇丰，学到了很多应用数学的理论知识，掌握了用数学建模解决实际问题的基本方法，也锻炼了编程实践的能力。在以后的课程中，希望助教能对每次的实验进行详细的解答，希望老师在课堂上介绍更多数学应用的前沿知识。

4 附录：程序代码

4.1 Chap13-Ex7

```
data = [
    1 44.6 44 89.5 6.82 62 178
    2 45.3 40 75.1 6.04 62 185
    3 54.3 44 85.8 5.19 45 156
    4 59.6 42 68.2 4.90 40 166
    5 49.9 38 89.0 5.53 55 178
    6 44.8 47 77.5 6.98 58 176
    7 45.7 40 76.0 7.17 70 176
    8 49.1 43 81.2 6.51 64 162
```

```
    9 39.4 44 81.4 7.85 63 174
    10 60.1 38 81.9 5.18 48 170
    11 50.5 44 73.0 6.08 45 168
    12 37.4 45 87.7 8.42 56 186
    13 44.8 45 66.5 6.67 51 176
    14 47.2 47 79.2 6.36 47 162
    15 51.9 54 83.1 6.20 50 166
    16 49.2 49 81.4 5.37 44 180
    17 40.9 51 69.6 6.57 57 168
    18 46.7 51 77.9 6.00 48 162
    19 46.8 48 91.6 6.15 48 162
    20 50.4 47 73.4 6.05 67 168
    21 39.4 57 73.4 7.58 58 174
    22 46.1 54 79.4 6.70 62 156
    23 45.4 52 76.3 5.78 48 164
    24 54.7 50 70.9 5.35 48 146
];

y = data(:,2);
xs = data(:,3:7);
x1 = xs(:,1); x2 = xs(:,2); x3 = xs(:,3); x4 = xs(:,4); x5 = xs(:,5);

% plot scatter
if 0
    for i = 1:5
        figure; scatter(xs(:,i),y,'+'); grid on;
        xlabel(['x' int2str(i)]); ylabel('y');
    end
end

% single param
if 0
    X = [ones(size(y)) x3];
    [b,bint,r,rint,stats] = regress(y,X);
    figure; rcoplot(r,rint); xlabel('Case Number'); ylabel('Residuals
        '); title('');

    x = [4 9];
    yHat = b(1) + b(2) * x;
    figure; plot(x,yHat); hold on; scatter(x3,y,'+'); grid on; xlabel
        ('x'); ylabel('y');

    polytool(x3,y);
end

```

```
% two param
if 0
    stepwise(xs, y);
    X = [ones(size(y)) x1 x3];
    [b,bint,r,rint,stats] = regress(y,X);
    figure; rcoplot(r,rint); xlabel('Case Number'); ylabel('Residuals'); title('');
end

% multi param
if 0
    stepwise(xs, y);
    X = [ones(size(y)) x1 x3 x5];
    [b,bint,r,rint,stats] = regress(y,X);
end

% outlier
if 1
    X = [ones(size(y)) x1 x3 x5];
    mask = true(size(y));
    mask(4) = 0; mask(10) = 0; mask(15) = 0; mask(17) = 0; mask(23) = 0;
    X = X(mask,:);
    y = y(mask);
    [b,bint,r,rint,stats] = regress(y,X);
    figure; rcoplot(r,rint); xlabel('Case Number'); ylabel('Residuals'); title('');
end
```

4.2 Chap13-Ex9

```
data = [
    1 6 28.1
    1 7 32.3
    1 8 34.8
    1 9 38.2
    1 10 43.5
    2 6 65.3
    2 7 67.7
    2 8 69.4
    2 9 72.2
    2 10 76.9
    3 6 82.2
    3 7 85.3
```

```
    3 8 88.1
    3 9 90.7
    3 10 93.6
];
x1 = data(:,1); x2 = data(:,2); y = data(:,3);

% x1 as a normal param
if 0
    X = [x1 x2];
    X1 = [ones(size(y)) X];
    stepwise(X, y);
    [b,bint,r,rint,stats] = regress(y,X1);
    figure; rcoplot(r,rint); xlabel('Case Number'); ylabel('Residuals'); title('');
end

% use two param
if 0
    x1b = double(x1 == 2);
    x1a = double(x1 == 3);
    X = [x1a x1b x2];
    X1 = [ones(size(y)) X];
    stepwise(X,y);
    mask = true(size(y));
    mask(5) = 0;
    [b,bint,r,rint,stats] = regress(y(mask),X1(mask,:));
    figure; rcoplot(r,rint); xlabel('Case Number'); ylabel('Residuals'); title('');
end

% add interaction
if 1
    X = [x1a x1b x2 x1a.*x2 x1b.*x2];
    X1 = [ones(size(y)) X];
    stepwise(X,y);
    [b,bint,r,rint,stats] = regress(y,X1);
    figure; rcoplot(r,rint); xlabel('Case Number'); ylabel('Residuals'); title('');
end
```

4.3 Chap13-Ex13

```
data = [
    1981 0 43.65
```

```
    1982 1 109.86
    1983 2 187.21
    1984 3 312.67
    1985 4 496.58
    1986 5 707.65
    1987 6 960.25
    1988 7 1238.75
    1989 8 1560.00
    1990 9 1824.29
    1991 10 2199.00
    1992 11 2438.89
    1993 12 2737.71
];

t = data(:,2);
yt = data(:,3);

% logistic linear
L = 3000;
y = log(L ./ yt - 1);
t1 = [ones(size(yt)) t];
[b,bint,r,rint,stats] = regress(y,t1);
a = exp(b(1));
k = -b(2);
beta = [L a k];
yt_hat = logistic(beta,t);
mse = sum((yt_hat - yt).^2) / (length(yt) - 2);

% logistic nonlinear
beta0 = [L a k];
[beta,R,J,CovB,MSE,ErrorModelInfo] = nlinfit(t, yt, @logistic, beta0)
    ;

% gompertz nonlinear
L = 3000; b = 30; k = 0.4;
beta0 = [L b k];
[beta,R,J,CovB,MSE,ErrorModelInfo] = nlinfit(t, yt, @gompertz, beta0)
    ;

% plot
t_plt = 0:0.01:12;
yt_plt = gompertz(beta, t_plt);
figure; plot(t_plt, yt_plt); hold on; scatter(t, yt);
grid on; xlabel('t'); ylabel('y_t');

```

```
function yt = logistic(beta,t)
    L = beta(1); a = beta(2); k = beta(3);
    yt = L ./ (1 + a * exp(-k .* t));
end

function yt = gompertz(beta,t)
    L = beta(1); b = beta(2); k = beta(3);
    yt = L * exp(-b * exp(-k * t));
end
```