



高中数学建模实践研究

河源中学 李艾 蔡维

2025年6月

目录

一、新手如何备赛

二、数学建模案例分析

三、建模常用软件介绍

四、优秀论文解析

五、答疑环节



一、新手如何备赛

明确团队分工

建模手

熟练掌握所有模型（如优化模型、评价模型、预测模型、统计模型等）的用处和特点

编程手

根据模型产出论文中所需要的数据及图表，也就是主攻算法，最好是熟悉Matlab，Python语言

论文手

掌握LaTeX、Word这两个常用排版软件，可以先上知网阅读一些文献，先按照数学建模常用的模板去梳理论文思路



二、数学建模案例分析

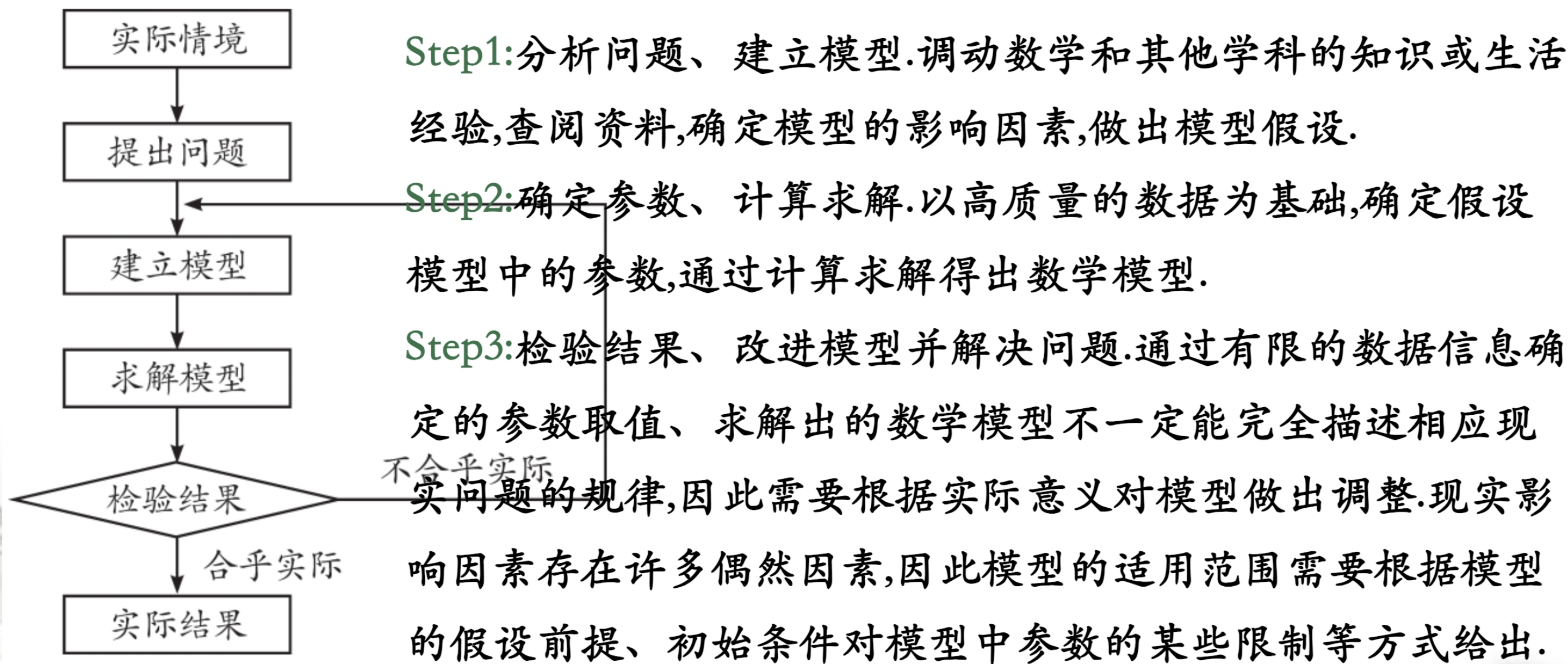


图2 数学建模基本流程图



二、数学建模案例分析

问题背景及变量分析

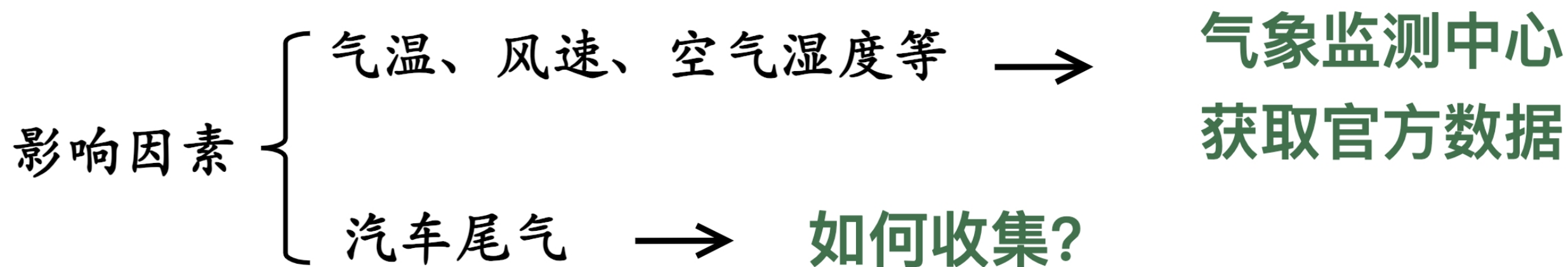
大气污染物 $PM_{2.5}$ （大气中直径小于或等于 $25\mu m$ 的颗粒物）的浓度超过一定限度会影响人的身体健康。**如何研究 $PM_{2.5}$ 的浓度是否受到汽车尾气、气候状况等因素的影响。**

1. 变量分析

- 自变量（解释变量）：
 - 汽车流量（辆/24小时）
 - 平均气温（ $^{\circ}C$ ）
 - 风速（m/s）
 - 空气湿度（ g/m^3 ）
- 因变量（响应变量）：
 - $PM_{2.5}$ 的平均浓度（ $\mu g/m^3$ ）



二、数学建模案例分析



模型假设 (model assumption) : 假设每一辆汽车排出的尾气量是相等的, 可以根据监测所得汽车流量来估计尾气排放量



二、数学建模案例分析

大气污染物 $PM_{2.5}$ （大气中直径小于或等于 $25\mu m$ 的颗粒物）的浓度超过一定限度会影响人的身体健康。研究 $PM_{2.5}$ 的浓度是否受到汽车尾气、气候状况等因素的影响。研究人员选择了24个社会经济发展水平相近的城市，在每个城市选择一个交通点建立监测点，统计每个监测点24小时内过往的汽车流量（单位：辆），同时在低空相同的高度测定每个监测点该时间段的平均气温（单位： $^{\circ}C$ ）、风速（单位： m/s ）、空气湿度（绝对湿度单位： g/m^3 ）以及空气中 $PM_{2.5}$ 的平均浓度（单位： $\mu g/m^3$ ），得到如下表所示数据：



二、数学建模案例分析

收集数据

表1 24个城市交通点空气中PM_{2.5}浓度监测数据

城市 编号	汽车 流量 /辆	平均 气温 /℃	空气 湿度/ (g·m ⁻³)	风速/ (m·s ⁻¹)	PM _{2.5} 浓度/ (μg·m ⁻³)	城市 编号	汽车 流量/辆	平均 气温/℃	空气 湿度/ (g·m ⁻³)	风速/ (m·s ⁻¹)	PM _{2.5} 浓度/ (μg·m ⁻³)
1	1300	20.00	80	0.45	66	13	950	22.50	69	2.00	35
2	1440	23.00	57	0.50	76	14	1440	21.50	79	2.40	58
3	790	26.50	64	1.50	21	15	1080	28.50	59	3.00	29
4	1650	23.00	84	0.40	170	16	1840	26.00	73	1.00	140
5	1760	29.50	72	0.90	156	17	1120	35.00	92	2.80	43
6	1750	30.00	76	0.80	120	18	1660	20.00	83	1.45	69
7	1200	22.50	69	1.80	72	19	1540	23.00	57	1.50	87
8	1500	21.80	77	0.60	120	20	960	24.80	67	1.50	45
9	1200	27.00	58	1.70	100	21	1780	23.30	83	0.90	222
10	1480	27.00	65	0.65	129	22	1500	27.00	65	0.65	145
11	1820	22.00	83	0.40	135	23	1060	26.00	58	1.83	34
12	1440	28.00	68	2.00	99	24	1440	28.00	68	2.00	105



二、数学建模案例分析

2. 确定分析方法

这是一个典型的**多元回归分析**问题，因为：

- 有一个连续的因变量（PM2.5浓度）。
- 有多个连续的自变量（汽车流量、气温、风速、湿度）。
- 目标是建立这些自变量与因变量之间的关系模型。



二、数学建模案例分析

3. 模型构建

3.1 多元线性回归模型

最基本的模型是多元线性回归模型：

$$\text{PM}_{2.5} = \beta_0 + \beta_1 \cdot \text{汽车流量} + \beta_2 \cdot \text{气温} + \beta_3 \cdot \text{风速} + \beta_4 \cdot \text{湿度} + \epsilon$$

其中：

- β_0 是截距。
- $\beta_1, \beta_2, \beta_3, \beta_4$ 是各自变量的回归系数。
- ϵ 是误差项。



二、数学建模案例分析

3.2 可能的模型扩展

- **非线性关系**：如果怀疑某些自变量与PM2.5的关系是非线性的（如气温与PM2.5可能存在U型关系），可以加入平方项或其他非线性项。
 - 例如： $\beta_2 \cdot \text{气温} + \beta_5 \cdot \text{气温}^2$
- **交互作用**：某些自变量可能共同影响PM2.5。
 - 例如：汽车流量和风速的交互作用（ $\beta_6 \cdot \text{汽车流量} \cdot \text{风速}$ ）。
- **标准化**：如果自变量量纲差异较大，可以对数据进行标准化（Z-score标准化）以改善模型稳定性。



二、数学建模案例分析

4. 数据预处理

在建模前，需要对数据进行以下检查和处理：

- **缺失值**：检查是否有缺失值，如果有，需要填补或删除。
- **异常值**：通过箱线图或Cook距离检查异常值，决定是否剔除。
- **正态性**：检查因变量（PM2.5）是否近似正态分布，必要时进行变换（如对数变换）。
- **相关性**：检查自变量之间的相关性，避免高度相关的自变量（如气温和湿度可能相关）。



二、数学建模案例分析

5. 模型拟合与评估

5.1 拟合模型

使用统计软件（如R、Python、SPSS）拟合多元线性回归模型。以R为例：

构建多元线性回归模型

```
model <- lm(pm_concentration ~ traffic_flow + avg_temperature + humidity + wind_speed,  
            data = org_data)
```

```
summary(model)
```

查看方差分析表

```
anova(model)
```

查看模型系数

```
coef(model)
```


二、数学建模案例分析

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-120.50343	69.16960	-1.742	0.0976 .
traffic_flow	0.13861	0.02318	5.980	9.37e-06 ***
avg_temperature	0.60727	1.82434	0.333	0.7429
humidity	0.10491	0.70768	0.148	0.8837
wind_speed	-0.17887	0.13622	-1.313	0.2048

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 30.63 on 19 degrees of freedom

Multiple R-squared: 0.7046, Adjusted R-squared: 0.6424

F-statistic: 11.33 on 4 and 19 DF, p-value: 7.172e-05

二、数学建模案例分析

##画出pm2.5和汽车流量的散点图，观察变量间的关系

```
plot(org_data$traffic_flow, org_data$pm_concentration,  
      xlab = "Traffic Flow (vehicles/hour)", # x轴标签  
      ylab = "PM2.5 Concentration ( $\mu\text{g}/\text{m}^3$ )", # y轴标签  
      main = "PM2.5 Concentration vs Traffic Flow", # 标题  
)
```

##画出pm2.5和平均温度的散点图

```
plot(org_data$avg_temperature, org_data$pm_concentration,  
      xlab = "avg_temperature ", # x轴标签  
      ylab = "PM2.5 Concentration ( $\mu\text{g}/\text{m}^3$ )", # y轴标签  
      main = "PM2.5 Concentration vs avg_temperature", # 标题  
)
```

##画出pm2.5和湿度的散点图

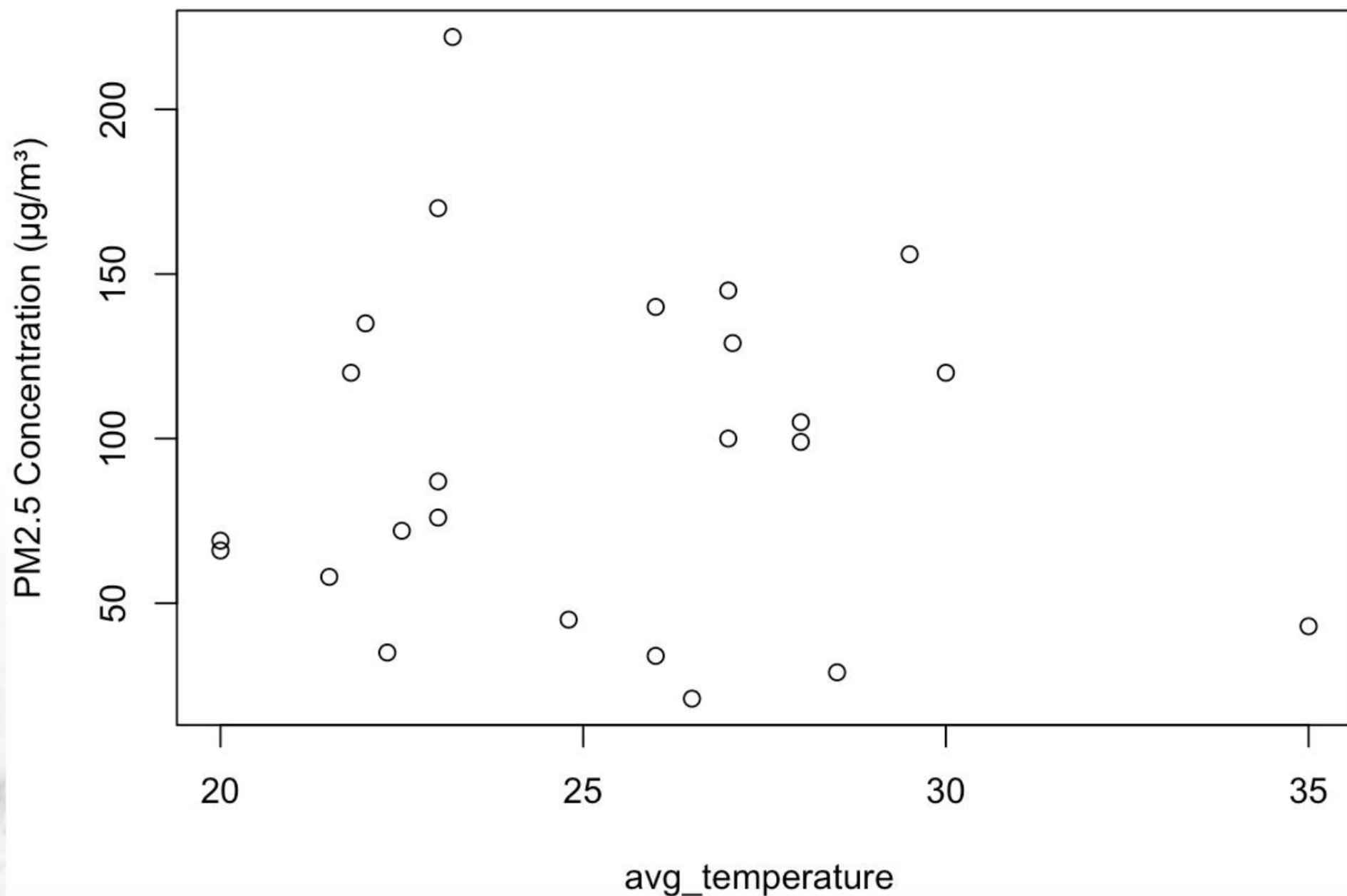
```
plot(org_data$humidity, org_data$pm_concentration,  
      xlab = "humidity", # x轴标签  
      ylab = "PM2.5 Concentration ( $\mu\text{g}/\text{m}^3$ )", # y轴标签  
      main = "PM2.5 Concentration vs humidity", # 标题  
)
```



二、数学建模案例分析

数据可视化

PM2.5 Concentration vs avg_temperature



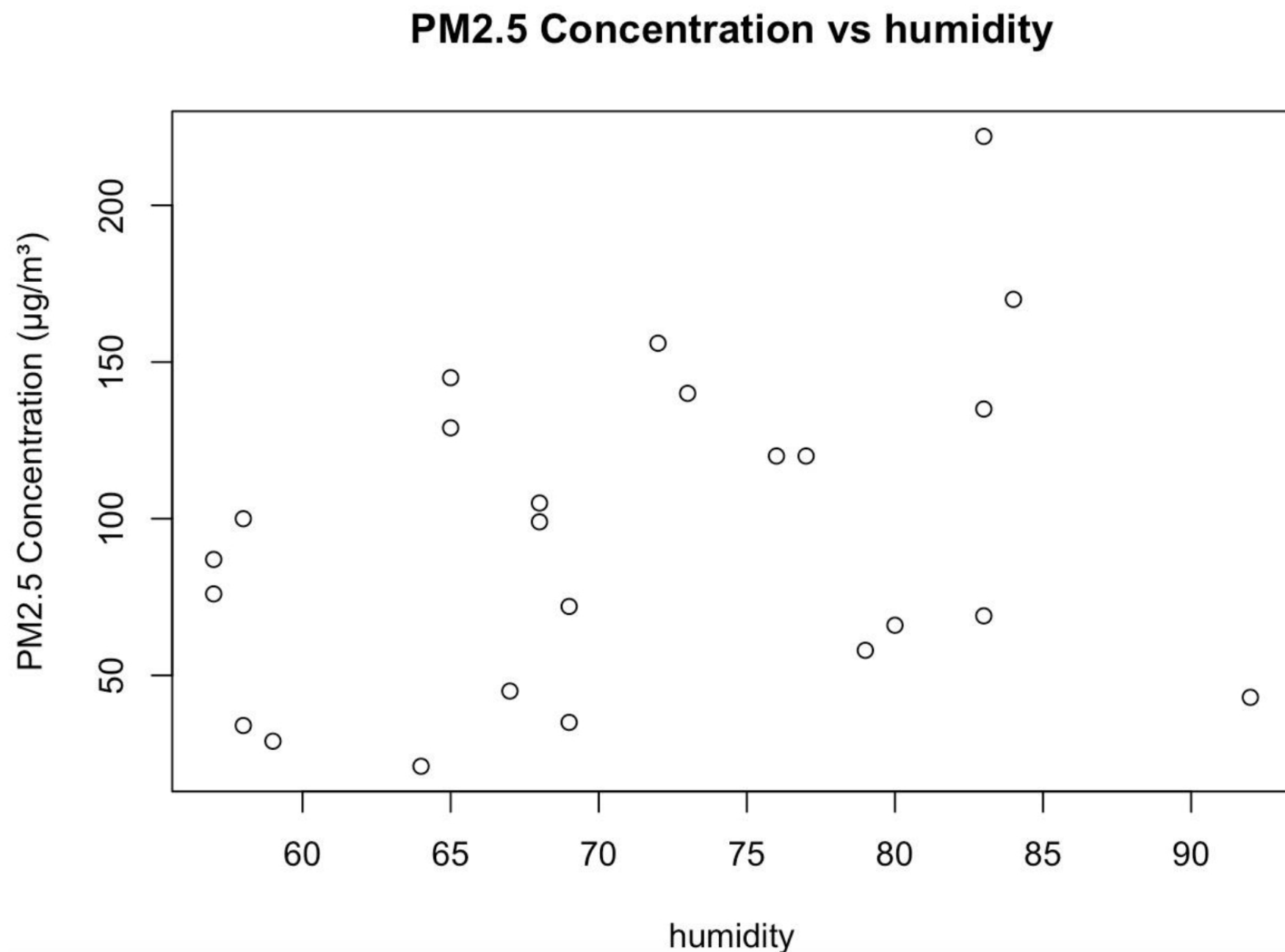
平均温度和pm2.5的散点图



河源市河源中学

二、数学建模案例分析

数据可视化



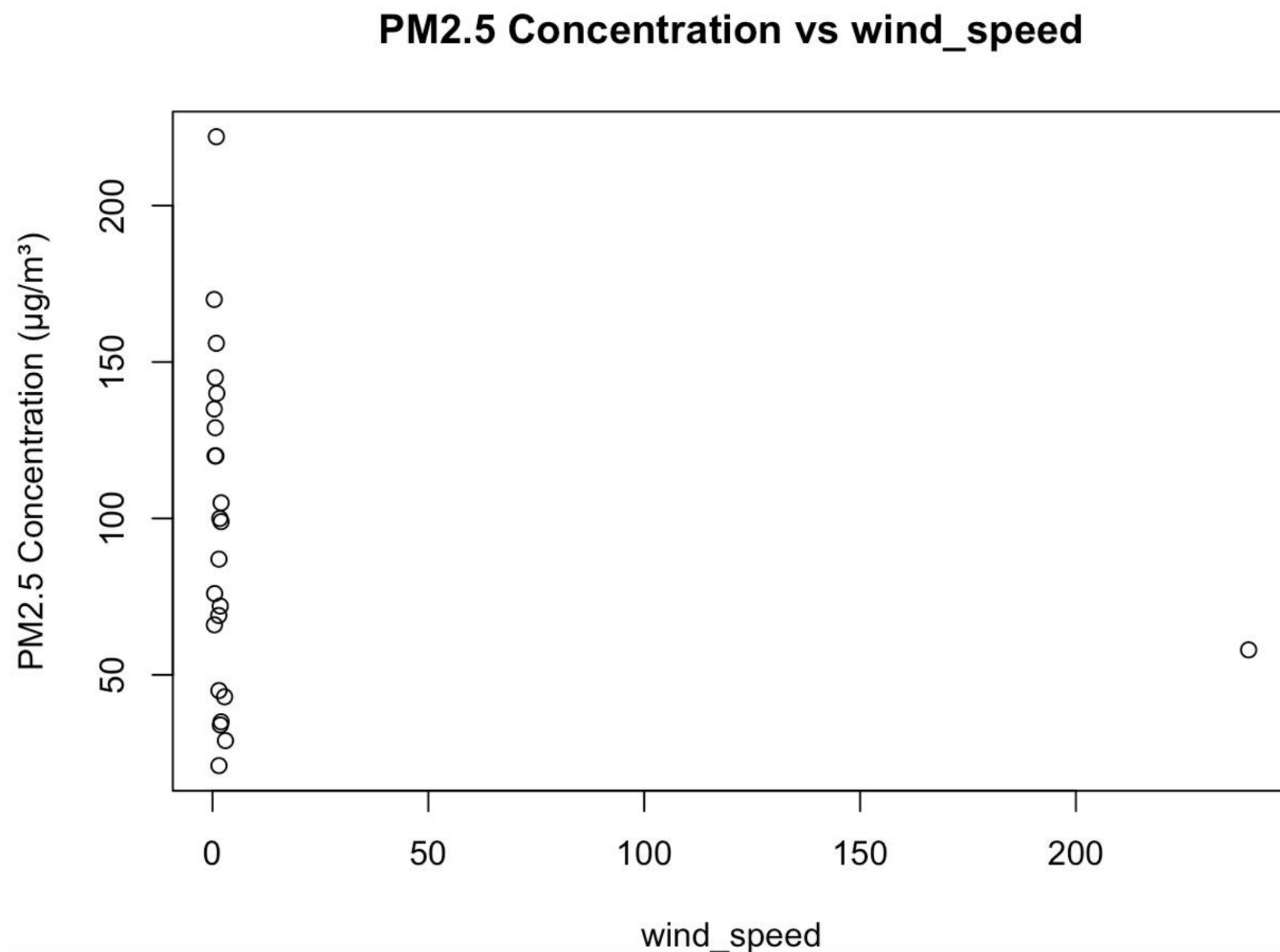
湿度和pm2.5的散点图



河源市河源中学

二、数学建模案例分析

数据可视化



风速与pm2.5的散点图



河源市河源中学

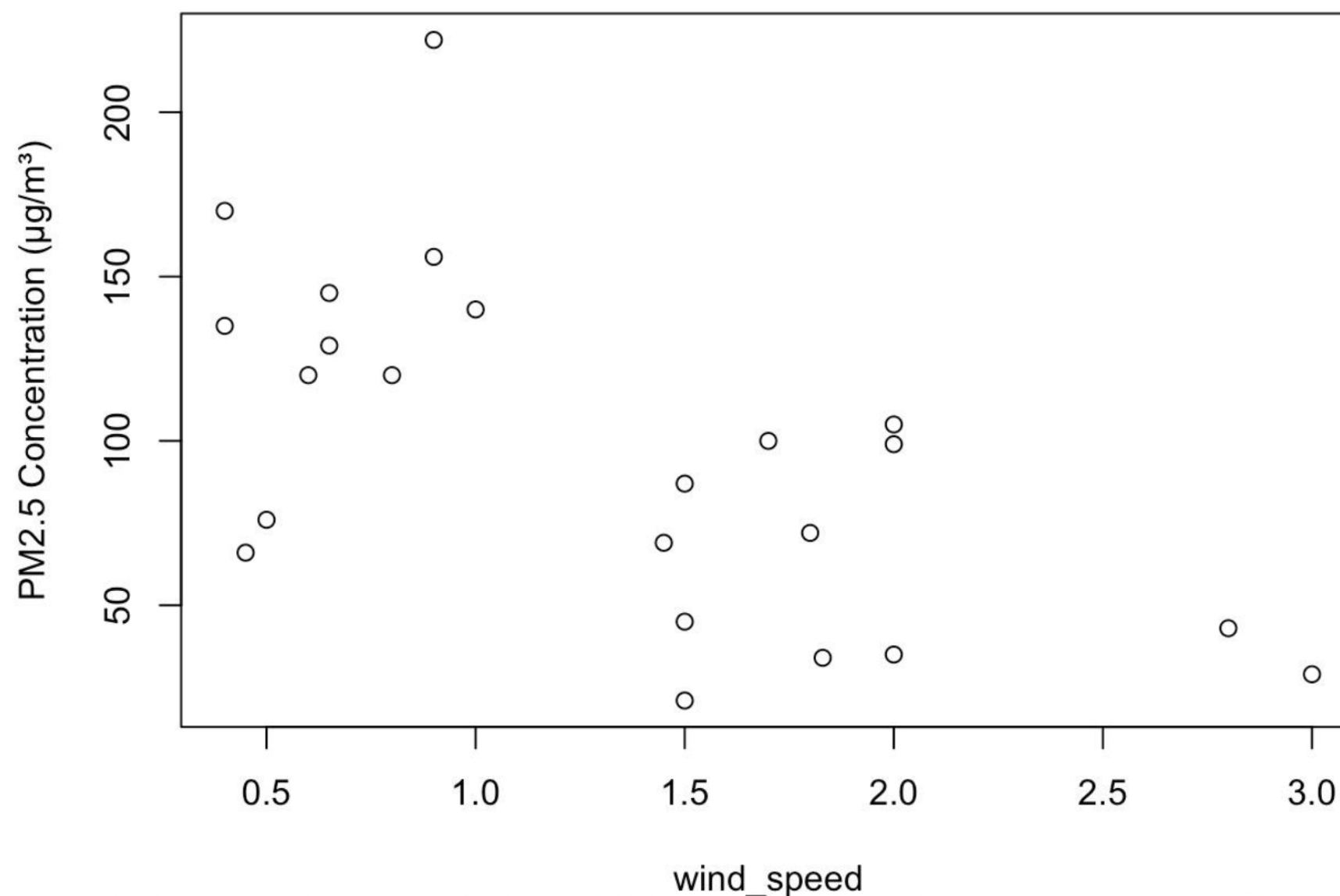
二、数学建模案例分析

```
#删除wind_speed超过200的异常行,画出pm2.5浓度和风速的散点图
data_clean <- org_data %>%
  filter(wind_speed <= 200)
plot(data_clean$wind_speed, data_clean$pm_concentration,
      xlab = "wind_speed", # x轴标签
      ylab = "PM2.5 Concentration ( $\mu\text{g}/\text{m}^3$ )", # y轴标签
      main = "PM2.5 Concentration vs wind_speed", # 标题
```



二、数学建模案例分析

PM2.5 Concentration vs wind_speed

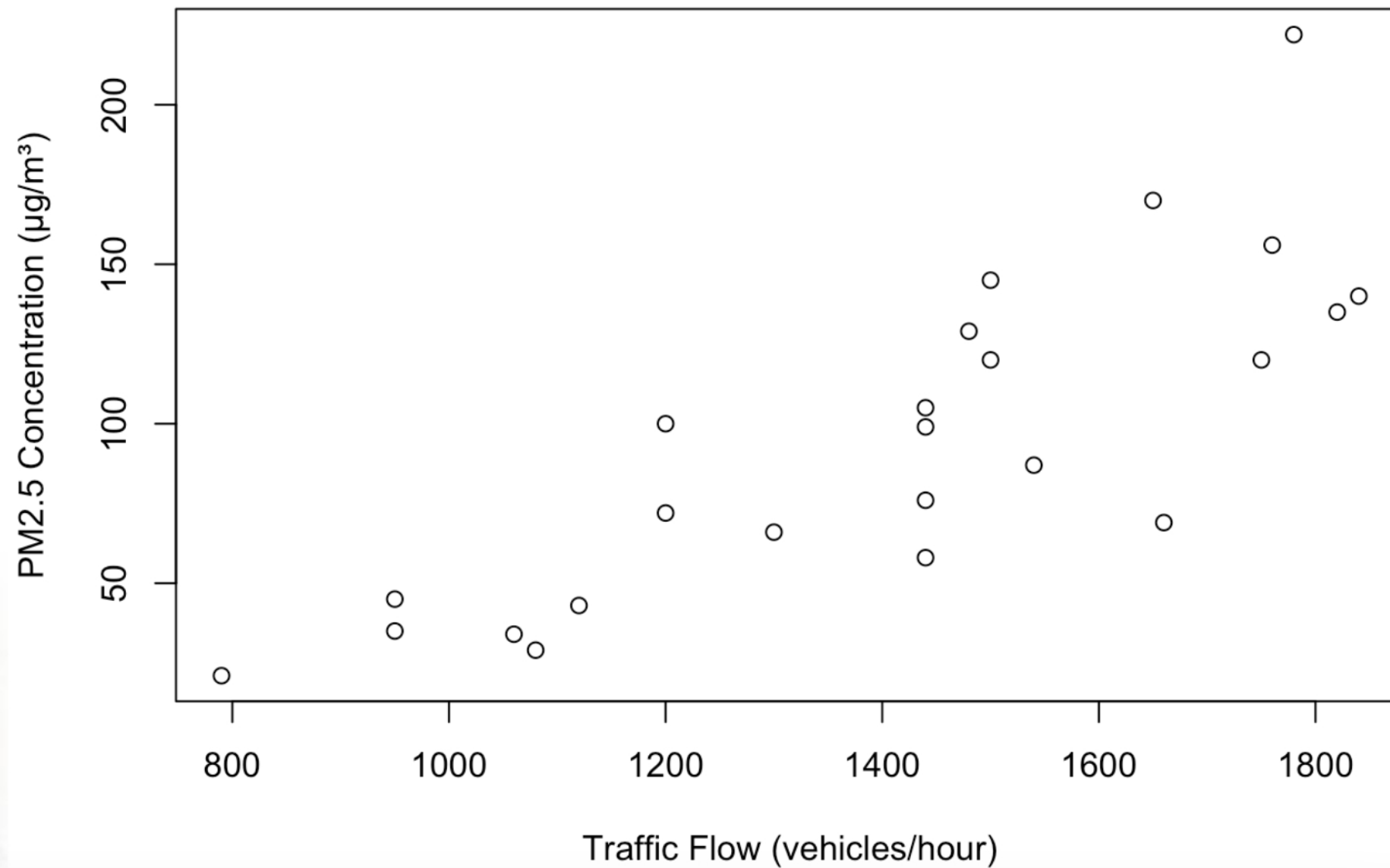


删除掉风速超过200的那一个城市的异常数据（第14号城市），风速和pm2.5的散点图

二、数学建模案例分析

数据可视化

PM2.5 Concentration vs Traffic Flow



汽车流量和pm2.5的散点图



二、数学建模案例分析

6.模型调整

提出假设H：汽车尾气排放与
 $PM_{2.5}$ 浓度之间存在显著相关性

```
##计算Pearson相关系数
cor_pearson <- cor(org_data$traffic_flow,
                    org_data$pm_concentration,
                    method = "pearson")
print(paste("Pearson相关系数:", round(cor_pearson, 3)))
##相关性检验（显著性p值）
cor_test <- cor.test(org_data$traffic_flow,
                     org_data$pm_concentration,
                     method = "pearson")

print(cor_test)
```


二、数学建模案例分析

Pearson's product-moment correlation

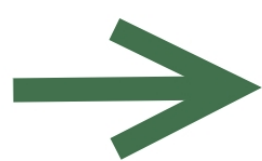
```
data: org_data$traffic_flow and org_data$pm_concentration  
t = 6.7063, df = 22, p-value = 9.669e-07  
alternative hypothesis: true correlation is not equal to 0  
95 percent confidence interval:  
 0.6215215 0.9190503  
sample estimates:  
 cor  
0.8194627
```

相关系数输出结果，相关系数
为0.8195，p值<0.05

结论： PM_{2.5}浓度随着汽车流量的增加，呈线性增长趋势，PM_{2.5}浓度与汽车流量之间是一种相关关系；

二、数学建模案例分析

PM_{2.5}浓度与汽车流量之间是一种相关关系



计算汽车流量与PM_{2.5}浓度两个变量之间的样本相关系数 (correlation coefficient)

1. 相关系数 (corr) : 范围在 $[-1,1]$ 之间, corr越接近于1说明两个变量之间的正相关性越强, 相关系数越接近于-1, 说明两个变量具有强负相关性。corr越接近0, 说明两个变量相关关系越弱/无线性相关关系。

2. 假设检验: 原假设 H_0 : 两个变量无相关性, 计算p值 (p-value) 。

若p值<显著性水平 (eg: 0.05) , 则拒绝原假设, 认为两个变量之间有相关关系, 否则无法拒绝原假设。

二、数学建模案例分析

线性回归

构建一元线性回归模型 $\begin{cases} Y = bx + a + e \\ E(e) = 0, D(e) = \sigma^2 \end{cases}$

$$\hat{y} = \hat{b}x + \hat{a}, \text{ 其中}$$

$$\hat{b} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad \hat{a} = \bar{y} - \hat{b}\bar{x}$$

$$\hat{a} = -99.60$$

$$\hat{b} = 138.60$$

一元线性回归方程:

$$\hat{y} = 138.60x - 99.60$$

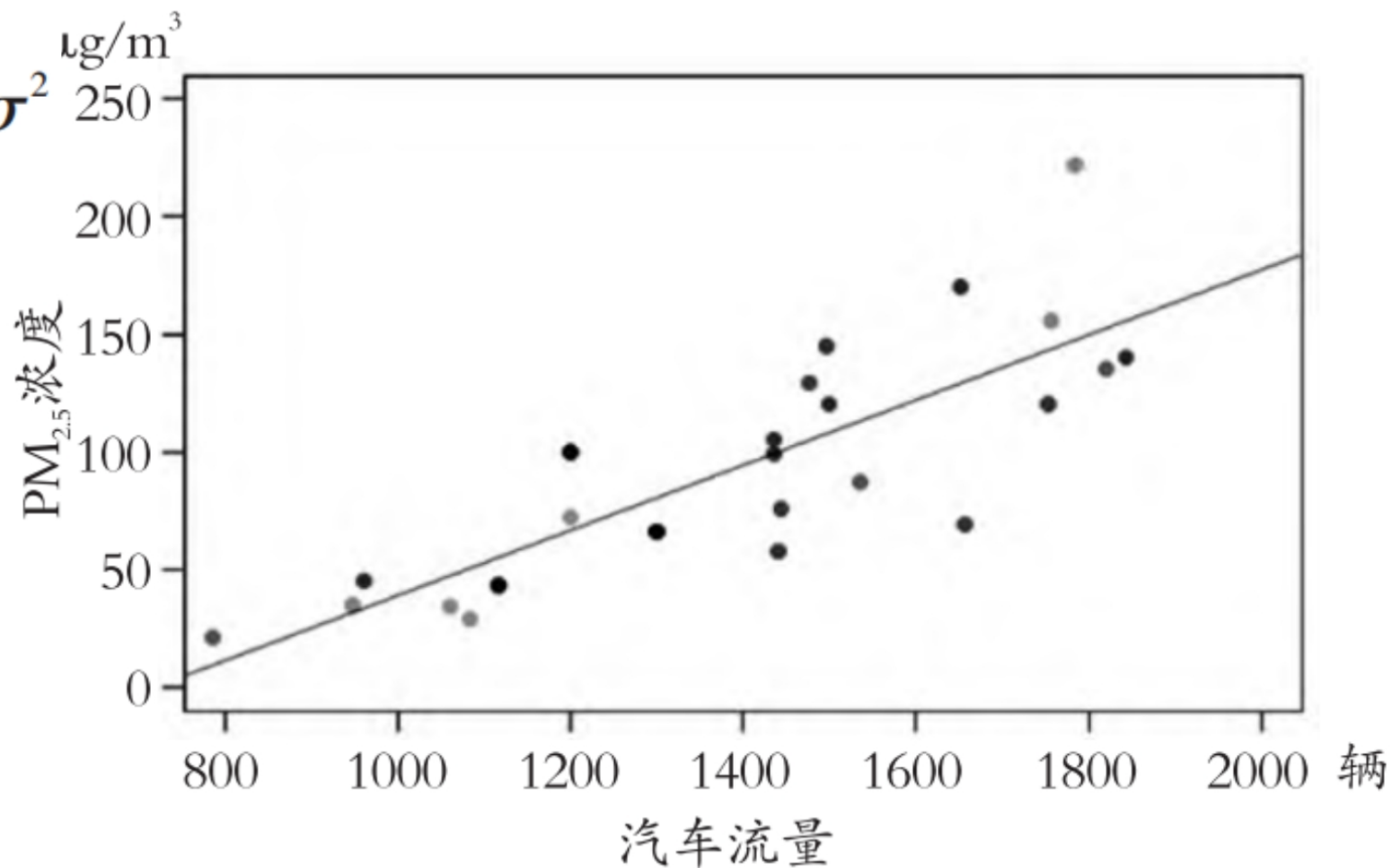


图4 汽车流量与PM_{2.5}浓度相关性统计模型

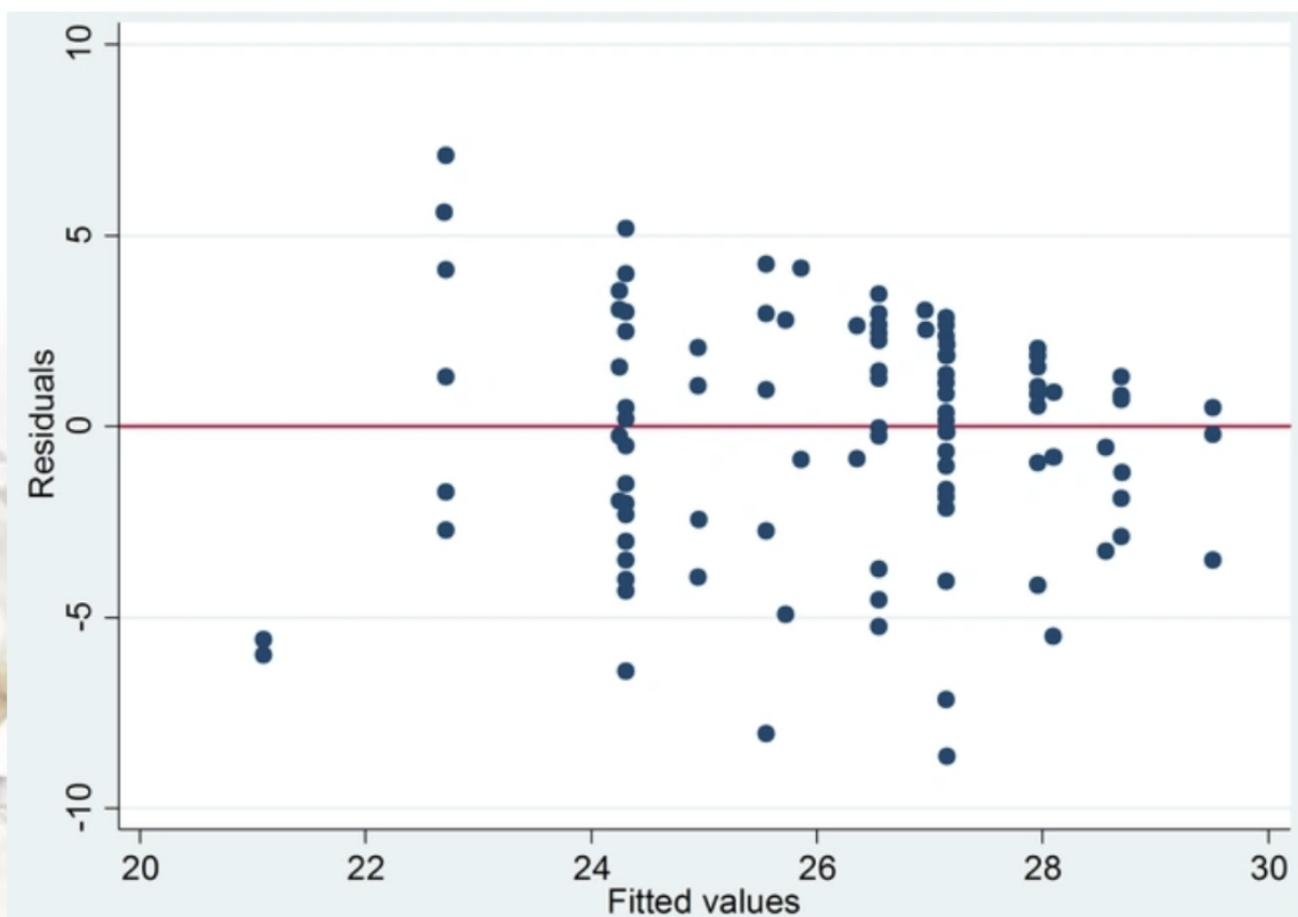
二、数学建模案例分析

模型检验

残差 (Residual) 分析：残差是观测值 y_i 与模型预测值 $\hat{y}_i$ 的差值

$$e_i = y_i - \hat{y}_i$$

在回归分析中，残差分析是验证模型假设（如线性、独立性、正态性）和识别模型缺陷的关键步骤。



- 图形法：直观快速（残差图、Q-Q图）
- 统计检验：定量验证（正态性、异方差性、自相关）
- 异常点检测：杠杆值、Cook距离
- 模型改进：根据问题调整模型（如变换变量、添加交互项）

通过系统检查残差，可以确保回归模型满足基本假设，提高结果的可靠性。

二、数学建模案例分析

模型优化

PM2.5的来源和影响因素

1. **汽车流量**：研究表明，汽车流量与PM2.5浓度之间存在线性关系。汽车流量每增加1千辆，PM2.5的平均浓度增加 $127.9 \mu\text{g}/\text{m}^3$ 3。
2. **工业排放**：工业活动也是PM2.5的重要来源之一，包括燃煤、冶金等过程产生的废气排放 1。
3. **城市规划和管理**：城市规划和管理措施，如绿化、交通管理、工业布局等，对减少PM2.5浓度有显著影响 1。



二、数学建模案例分析

模型优化

模型局限性	优化方法
原始数据存在异常值和离群点	学生化残差检验/删除异常值/对数化处理/中位数、分位数填充
模型的变量之间存在非线性关系	选择非线性模型（non-linear model）
变量之间存在多重共线性	正则化处理（岭回归/Lasso回归）
模型过拟合（overfitting）	增加样本数据量/降低模型复杂度/交叉验证

二、数学建模案例分析

进行统计解释和应用

学生根据回归模型的结果进行统计解释，并结合实际问题提出建议：

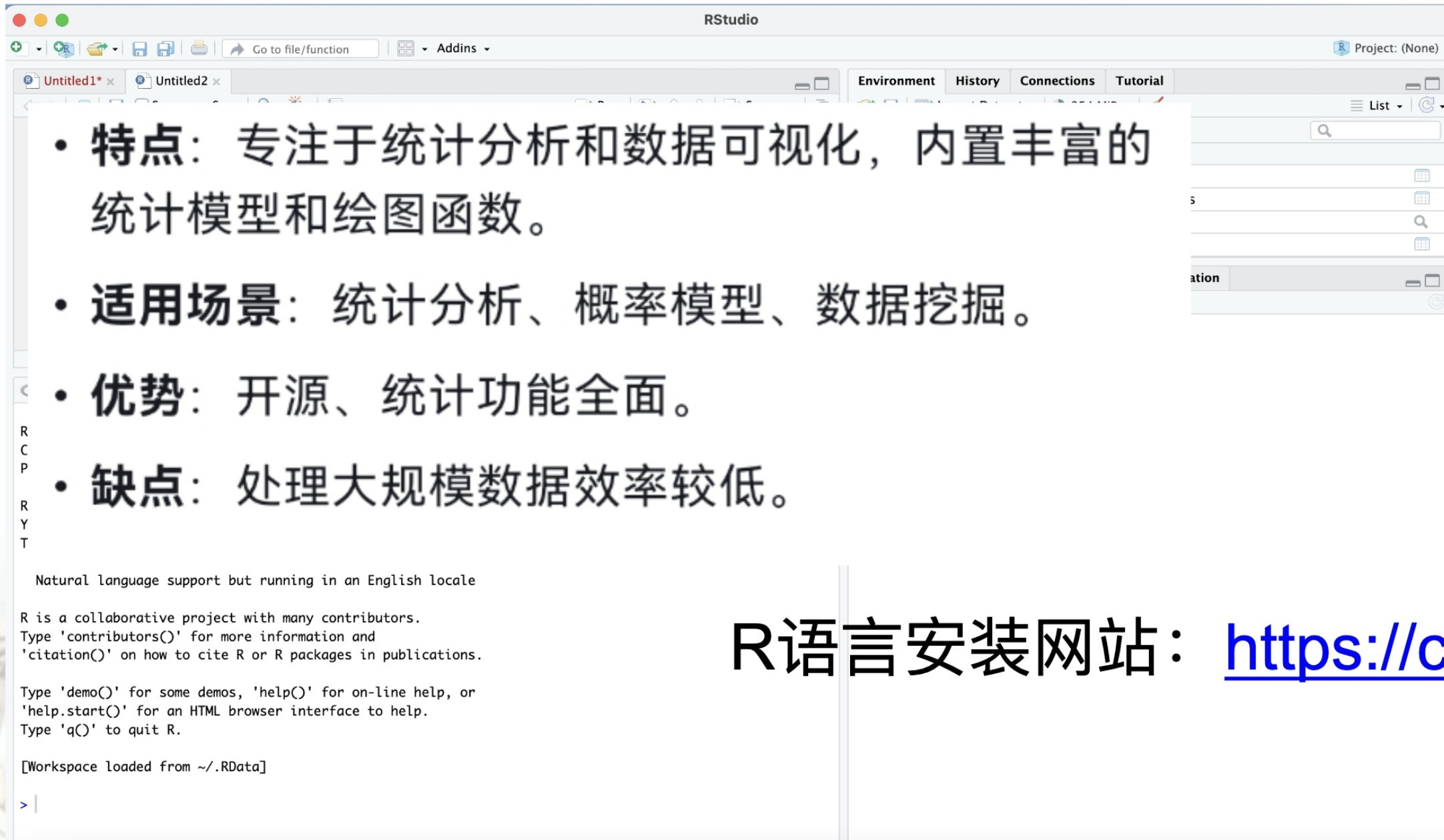
- ①我国规定空气中 $\text{PM}_{2.5}$ 浓度的安全标准为年平均浓度 $35\mu\text{g}/\text{m}^3$ ，24小时平均浓度 $75\mu\text{g}/\text{m}^3$ ，那么当汽车流量为1300辆和2300辆时，计算 $\text{PM}_{2.5}$ 的浓度并判断是否需要采取预警和干预措施？
- ②某城市为使 $\text{PM}_{2.5}$ 浓度平均值在 $60—120\mu\text{g}/\text{m}^3$ ，拟对汽车流量作适当控制，请用所得的回归模型，计算该城市24小时的汽车流量的取值范围，并提出政策建议。

根据问题，学生能够利用多元回归模型，根据特定的汽车流量预测 $\text{PM}_{2.5}$ 浓度的变化趋势，为交通管理或环境保护政策提供数据支持。



三、建模常用软件介绍

R语言



The screenshot shows the RStudio application window. The top menu bar includes 'File', 'Edit', 'Session', 'View', 'Help', and 'Addins'. The toolbar contains icons for file operations and running code. The main editor area has two tabs: 'Untitled1*' and 'Untitled2'. The right-hand side has four panes: 'Environment', 'History', 'Connections', and 'Tutorial'. The 'Environment' pane shows a list of objects. The 'History' pane shows a list of executed commands. The 'Connections' pane shows a list of connections. The 'Tutorial' pane shows a list of tutorials. The console at the bottom displays the following text:

```
Natural language support but running in an English locale  
  
R is a collaborative project with many contributors.  
Type 'contributors()' for more information and  
'citation()' on how to cite R or R packages in publications.  
  
Type 'demo()' for some demos, 'help()' for on-line help, or  
'help.start()' for an HTML browser interface to help.  
Type 'q()' to quit R.  
  
[Workspace loaded from ~/.RData]  
> |
```

- **特点：**专注于统计分析和数据可视化，内置丰富的统计模型和绘图函数。
- **适用场景：**统计分析、概率模型、数据挖掘。
- **优势：**开源、统计功能全面。
- **缺点：**处理大规模数据效率较低。



R语言安装网站：<https://cran.r-project.org>



河源市河源中学

三、建模常用软件介绍

Python



Donate



Search

GO

Socialize

ries

News

Events

ation

with Python, and expression syntax is
ators +, -, * and / work as
can be used for grouping. [More](#)
[ons in Python 3.](#)

3

4

5

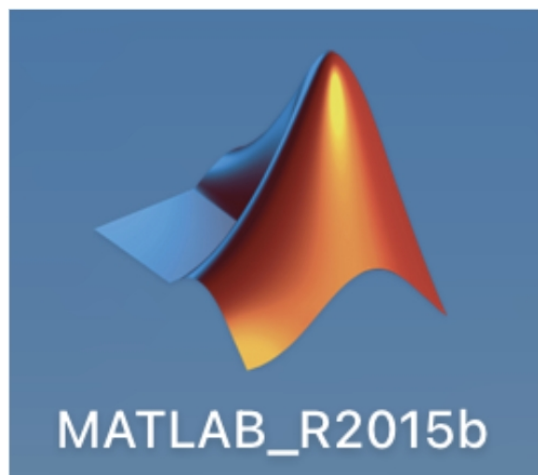
- **特点：** 开源生态，通过库（如NumPy、SciPy、SymPy）实现科学计算，Matplotlib/Seaborn用于可视化。
- **适用场景：** 数据分析、机器学习、符号计算、优化问题。
- **优势：** 免费、灵活性高，适合复杂算法开发。
- **缺点：** 需要编程基础，调试可能较复杂。

Python is a programming language that lets you work quickly
and integrate systems more effectively. >>> [Learn More](#)



河源市河源中学

三、建模常用软件介绍



MATLAB

- **特点：**强大的数值计算、矩阵运算和算法开发工具，内置大量数学函数和工具箱（如优化、统计、信号处理）。
- **适用场景：**工程仿真、数值分析、动态系统建模。
- **优势：**语法简单，可视化能力强，社区支持丰富。
- **缺点：**商业软件，价格较高。

matlab矩阵实验3

SPSS

- **特点：**图形化界面，适合非编程用户。
- **适用场景：**问卷调查、统计分析。
- **缺点：**高级功能需付费。

SPSS社会科学统计软件包



 **首页**

专业开发者社区CSDN



河源市河源中学

三、建模常用软件介绍

CSDN 博客 下载 学习 社区 GitCode InsCodeAI 会议

登录

python 马尔萨斯模型

C一下

全站

知道 R1

博客

下载

课程

用户

专栏

更多

VIP买1年送1年

综合

最新

热门

☐ VIP内容

开发者首选的AI搜索引擎

python 马尔萨斯模型

深度思考



人口预测模型-指数模型、logistic模型、马尔萨斯模型、Leslie模型介绍, pyth...

本文将详细介绍四种常用的人口预测模型: 指数模型、logistic模型、马尔萨斯模型和Leslie模型, 并提供Python和MATLAB的代码实现, 帮助...

1824 1 721KB

zcongfly 2023-01-21

数学建模之Python-微分方程建模实例

微分方程的数学求解用到再细学 见python数学建模与实验pdf 微分方程建模-Malthus模型 微分方程建模-Logistic模型

这里不细说了 感觉用不太到 也是调参预测人口数量的 微分方程建模-传染病模型 传染病模型例子: ...

3. 模型求解
考虑二阶常微分方程初值问题: $y'' + p(x)y' + q(x)y = r(x)$, $y(x_0) = y_0$, $y'(x_0) = y'_0$.
解法为:
 $y(x) = y_h(x) + y_p(x)$
其中 $y_h(x)$ 为齐次方程的通解, $y_p(x)$ 为非齐次方程的特解.

1981 3 1

Cabbage coder 2021-08-11

市河源中学

三、建模常用软件介绍

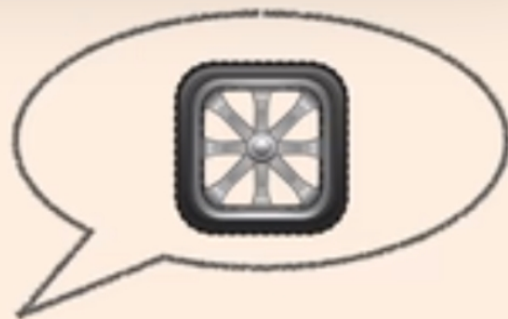
B站教程

3

你觉得这辈子都学不会编程？超基础Python课程，3小时快速入门 | 自学Python教程 | 零基础Python | 计算机二级Python | Python期末速成

479.0万 3.1万 2021-10-23 07:15:09 未经授权，禁止转载

一般都做不出来满分的作业



但写代码的时候



河源市河源中学

四、优秀论文解析

2

基于单目标优化模型的电梯空闲停靠问题



中国知网

摘要

本文主要研究了电梯的空闲停靠问题，根据生活中电梯的实际使用情况，利用函数思想与多元非线性方程组，建立数学模型，求解出一定情况下乘客平均等待时间最短的电梯停靠位置。

本文的模型考虑了不同的使用场景，客流情况，电梯数量等因素，利用易于理解的算法基于函数思想求出结果，并利用 matlab 对求解结果进行验证，具有高效性和准确性，为电梯的最佳默认停靠位置设计提供了参考依据。

关键词：单目标优化；电梯空闲停靠；构造函数思想



河源市河源中学

四、优秀论文解析

为了使求解过程更为直观，我们还可以用柱状图进行求解。

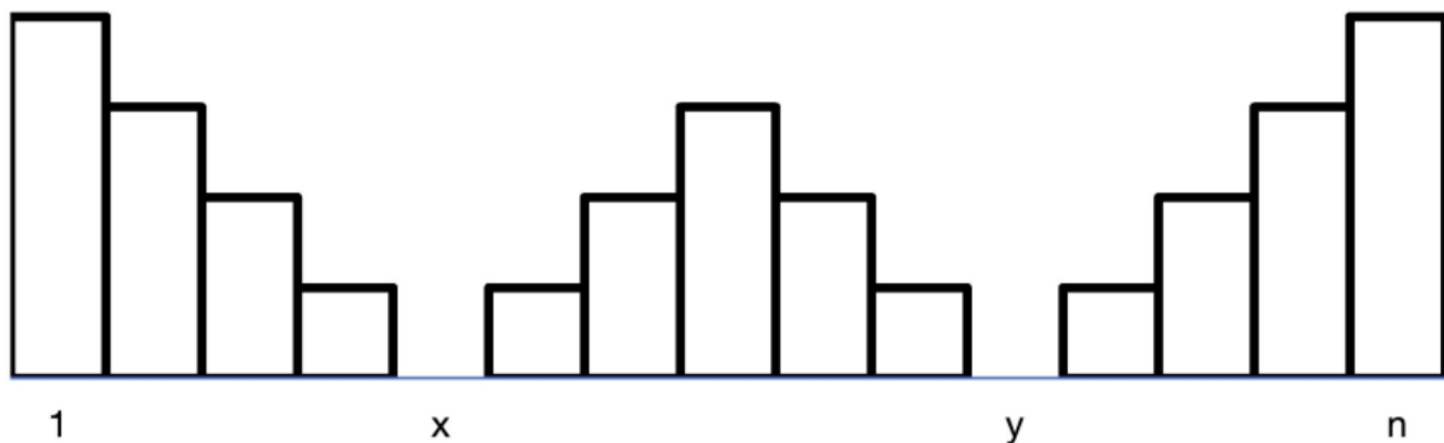


图1 住宅模式双电梯接客路程示意图

如图：横坐标表示楼层数，X，Y 表示两电梯分别停靠的楼层，各柱形的高度表示电梯接客移动的楼层数，面积表示电梯接客移动的路程。

为了便于表示，我们省略了大部分柱形(如 1 和 x 之间应有 $x-1$ 个柱形)，但不影响计算。

则所有柱形的面积之和减去第一个柱形的面积表示电梯除去 1 楼外接客的总路程，另有 $(x-1)(n-1)$ 表示电梯去 1 层接客的总路程。

通过计算面积我们可以得到乘客指数

$$E = \frac{3}{4}(x^2 + y^2) - \frac{1}{2}xy + (n - \frac{5}{2})x - \frac{1}{2}y + c$$

其余过程同上，通过柱形图，求解过程更为直观。

利用matlab绘制图示，工具使用熟练，结果可视化直观



四、优秀论文解析

5.2 “商场模式”的模型建立与求解

5.2.1 单电梯“商场模式”

6.2 模型评价

优点：在适当的假设下，可以较为简单直观地计算出不同情况下含参量的乘客指数，并以此求解给出最小值点，即电梯的最佳空闲停靠位置。

缺点：模型不够适用于电梯过于忙碌，没有空闲停靠机会的情况。

六. 模型验证与评价

6.2.1 模型的优点

模型复杂度低,乘客指数表达式的物理意义清晰；模型求解直观简单，结果简洁明了，求解算法有泛化性；运用 matlab 程序对求解结果进行验证，快捷准确。

6.2.2 模型的创新点

论文包含问题重述、模型假设、求解和验证等完整环节，符合数学建模基本流程



四、优秀论文解析

基于微分方程研究乒乓球侧旋轨迹与接法

摘要

本文主要研究了电梯的空闲停靠问题，根据生活中电梯的实际使用情况，利用函数思想与多元非线性方程组，建立数学模型，求解出一定情况下乘客平均等待时间最短的电梯停靠位置。

本文的模型考虑了不同的使用场景，客流情况，电梯数量等因素，利用易于理解的算法基于函数思想求出结果，并利用 matlab 对求解结果进行验证，具有高效性和准确性，为电梯的最佳默认停靠位置设计提供了参考依据。

关键词：单目标优化；电梯空闲停靠；构造函数思想

论文对题目背景的理解透彻，能将实际问题转化为明确的数学问题



四、优秀论文解析

5.1.2 问题一模型的求解

以乒乓桌的长度方向为纵轴 y ，宽度方向为横轴 x 建立坐标系，以球在空中某一点仅有纵向初速度为基准原点，考虑到纵向阻力（与纵向速度成正比），横向受玛格努斯力做的曲线运动，由物理学分析得到微分方程：

$$m \frac{dv_y}{dx} = -kv_y \quad ①$$

$$m \frac{dv_x}{dt} = F \quad ②$$

初始时， $t=0, v_x=0, v_y=v_0, x=y=0$

对②积分可得：

$$V_x = \frac{F}{m}t \quad ③$$

对①进行分离变量再进行积分可得：

$$v_y = e^{-\frac{k}{m}t + \ln v_0} \quad ④$$

由运动关系的：

$$\begin{aligned} x &= \int v_x dt \\ y &= \int v_y dt \end{aligned} \quad ⑤$$

可得到：

$$x = \frac{F}{2m}t^2 \quad ⑦$$

论文数学公式书写格式规范、准确，模型选择恰当、合理



四、优秀论文解析

五、模型建立与求解

数据的预处理

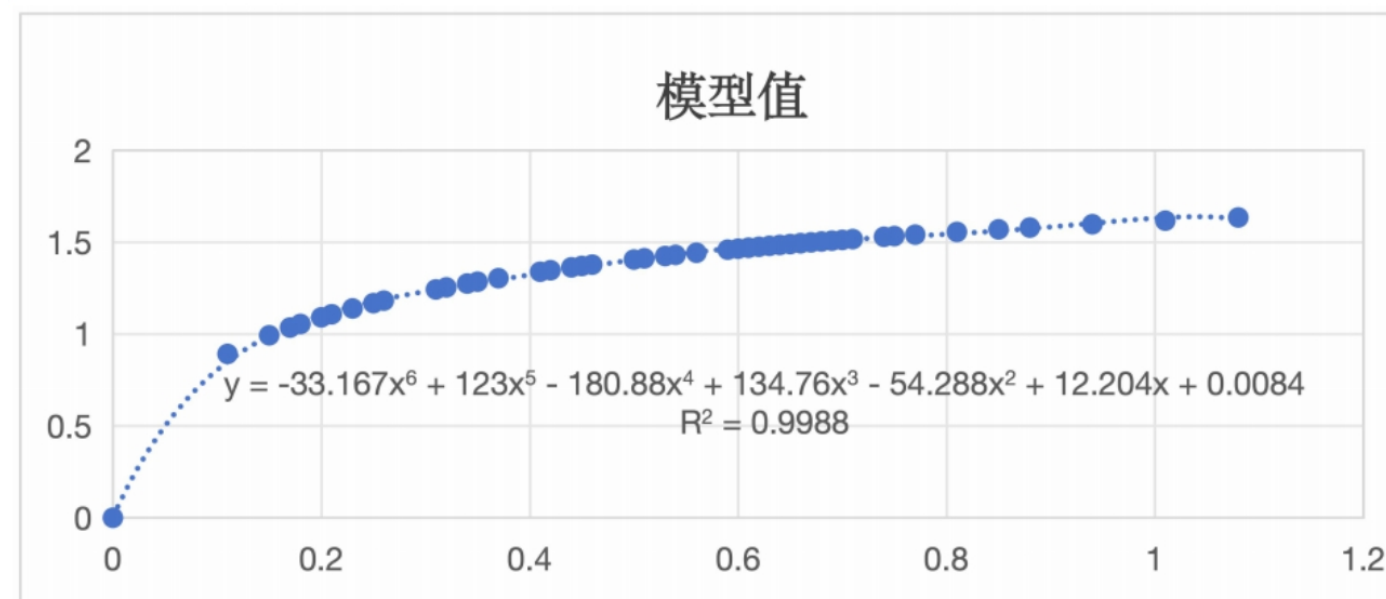
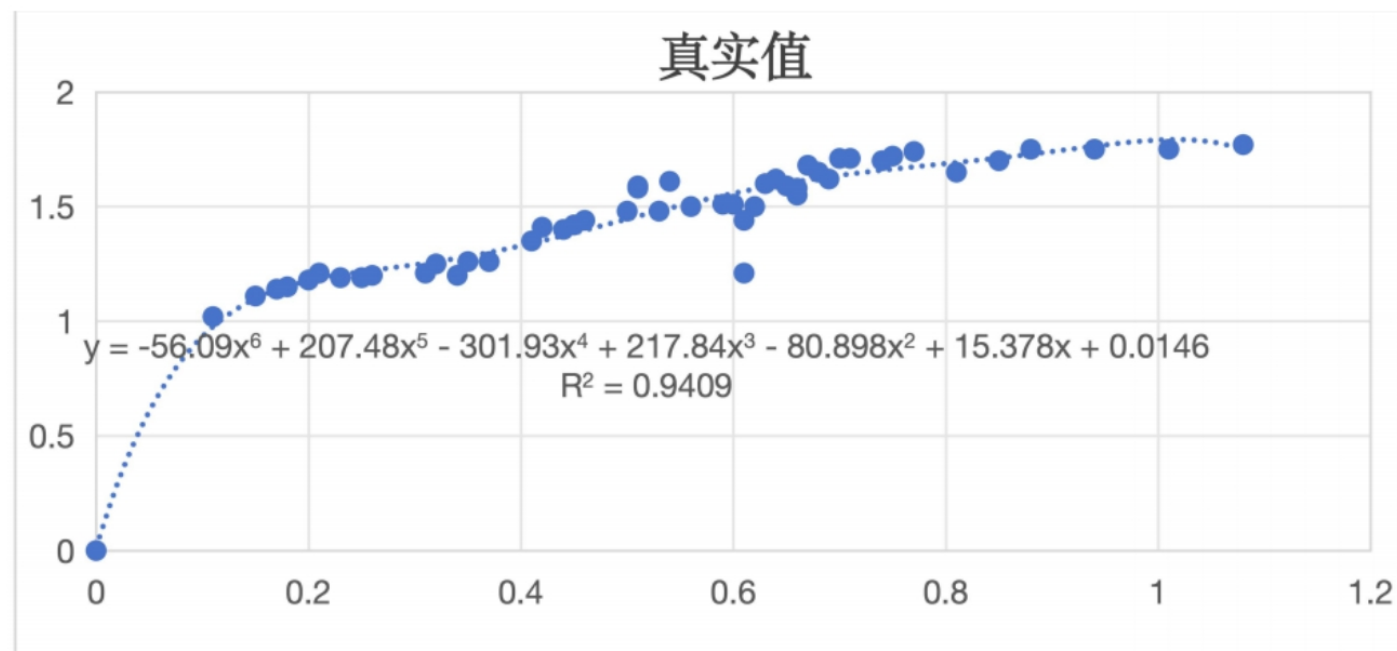
这里是把实验数据和所求得的模型数据进行了合并展示, 以便于后面的分析。
X, y 真实值是实验数据值, y 模型值是 x 真实值代入模型求得的数据。

x 真实值	y 真实值	x 真实值	y 模型值
0.61	1.21	0.61	1.47
0.66	1.55	0.66	1.49
0.50	1.48	0.50	1.41
0.59	1.51	0.59	1.46
0.69	1.62	0.69	1.51
0.70	1.71	0.70	1.51
0.65	1.59	0.65	1.49

数据来源真实可靠, 解释性强, 且包含数据预处理环节



四、优秀论文解析



计算模型值与真实值的平均相对误差:

$$\bar{E} \approx \frac{1}{50} \sum_{i=1}^{50} \left| \frac{y_i - \hat{y}_i}{y_i} \right| \approx 6\%$$

论文包含模型评价（误差分析、优缺点、优化），符合数学建模基本流程





河源市河源中学

五、答疑环节

2025年6月



河源市河源中学

