

# 建模竞赛及一个建模案例

2021年10月

# 一. 研究生数学建模竞赛

中国研究生数学建模竞赛作为教育部学位管理与研究生教育司、教育部学位与研究生教育发展中心指导、中国学位与研究生教育学会、中国科协青少年科技中心主办的“**中国研究生创新实践系列大赛**”主题赛事之一，是一项面向在校研究生进行数学建模应用研究的学术竞赛活动，是广大在校研究生**提高建立数学模型和运用互联网信息技术解决实际问题能力**，培养科研创新精神和团队合作意识的大平台。

2021年“华为杯”第十八届中国研究生数学建模竞赛由华南理工大学承办，由华为技术有限公司赞助

## 数学建模：培养综合素养

- 社会认可：社会、高校以及学生
- 数学建模：培养解决实际问题能力的理想环境、途径
  - 传统课程教学：知识学习、“浅”层次解决问题
  - 数学建模：更广范围的综合知识学习、灵活解决问题的技巧、深层次解决问题的方法
- 数学建模：培养分析和解决复杂问题的能力
  - 文献查阅（了解问题背景和解决当前问题的方法）
  - 数据的采集与分析处理（采集手段、处理方法）
  - 问题的抽象化、建模（直觉的培养，举一反三）
  - 算法设计与编程实现（可行、编程经验与技巧、算法时间空间复杂度）
- 综合体现：学习能力、实践能力、创新能力、协作能力、抗挫能力、写作能力

# 近几年全国大学生数学建模竞赛题目（问题+数学）

2012A题 葡萄酒的评价【统计问题】

2013A题 车道被占用对城市道路通行能力的影响

2014A题 嫦娥三号软着陆轨道设计与控制策略【控制-优化】

2015A题 太阳影子定位【视频数据分析-优化问题】

2016A题 系泊系统的设计【力学问题-优化】

2017A题 CT系统参数标定及成像【反问题】

2018A题 高温作业专用服装设计【偏微分方程反问题】

2019A题 高压油管的压力控制【微分方程反问题】

2020A题 炉温曲线【微分方程反问题】

**2021A题 “FAST”主动反射面的形状调节**

2012B题 太阳能小屋的设计【优化问题】

2013B题 碎纸片的拼接复原【匹配-优化问题】

2014B题 创意平板折叠桌【几何设计问题】

2015B题 “互联网+”时代的出租车资源配置【统计分析】

2016B题 小区开放对道路通行的影响【指标体系,统计分析】

2017B题 “拍照赚钱”的任务定价【影响因素, 统计分析】

2018B题 智能RGV的动态调度策略【优化问题】

2019B题 “同心协力”策略研究

2020B题 穿越沙漠【优化决策问题】

**2021B 题 乙醇偶合制备 C4 烯烃**

2019C题 机场的出租车问题

2020C题 中小微企业的信贷决策【决策问题】

**2021C题 生产企业原材料的订购与运输**

**变（绝对）：**背景领域（应用场景）

**不变（相对）：**建模思想、模型过程、求解的算法

2017 年中国研究生数学建模竞赛 B 题（华为公司命题）。

2017 年中国研究生数学建模竞赛 C 题

面向下一代光通信的 VCSEL 激光器仿真

航班恢复问题

2017 年中国研究生数学建模竞赛 D 题。

基于监控视频的前景目标提取。

2018年第十五届“华为杯”中国研究生数学建模竞赛(多无人机对组网雷达的协同干扰)-E题

第十五届“华为杯”中国研究生数学建模竞赛(对恐怖袭击事件记录数据的量化分析)—C题

第十五届“华为杯”中国研究生数学建模竞赛(航站楼扩增评估)—F题

第十五届“华为杯”中国研究生数学建模竞赛(卫星高度计海面高度异常资料获取潮汐调和常数方法及应用)—D题

第十五届“华为杯”中国研究生数学建模竞赛（光传送网建模与价值评估）—B题

第十五届“华为杯”中国研究生数学建模竞赛（跳台跳水体型系数设置的建模分析）—A题

PDF 2019年A题 无线智能传播模型.pdf

DOC 2019年B题 天文导航中的星图识别.doc

PDF 2019年C题 视觉情报信息分析.pdf

DOC 2019年D题 汽车行驶工况构建0914.docx

PDF 2019年E题 全球变暖? .pdf

PDF 2019年F题 多约束条件下飞行器航迹快速规划.pdf

DOC 2020年A题--华为题\_芯片相噪算法设计.docx

DOC 2020年B题--汽油辛烷值优化建模.docx

DOC 2020年C题--面向康复工程的脑电信号分析和判别模型.docx

DOC 2020年D题--无人机集群协同对抗.doc

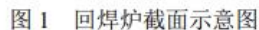
DOC 2020年E题--能见度估计与预测.doc

DOC 2020年F题--飞行器质心平衡供油策略优化.docx

- 专业背景更强  
(紧密结合科技前沿和实际应用)
- 研究问题更深入
- 模型、算法

## **二、建模竞赛论文写作**

## 2020A题 炉温曲线



- 摘要、印象: 20分
- 主要模型
- 主要算法
- 结果(合理、分析)

A 题 炉温曲线 评分参考, 仅供讨论用, 不作最后标准。

论文	摘要	问题一 (30)	问题二 (15)	问题三 (20)	问题四 (15)	印象	总分
	10	模型 算法 结果 15 7 8	模型 算法 结果 5 5 5	模型 算法 结果 5 10 5	模型 算法 结果 5 5 5	10	100

**问题一 (30分)**

**模型分 15分**

(1) 对问题进行分析建立传热机理模型 (热传导, 热对流, 热辐射, **模型简化**) (8-15分)。

(2) 无机理模型 (数据拟合等) (0-7分)

**算法分 7分**

(1) 能描述差分方程数值计算过程, 写清楚计算步骤 (5-7分)

**结果分 8分**

(1) 四温度值 138.07 (-10), 171.7 (+10), 183.32 (+10), 202.65 (+10) (2-6分)

(2) 有炉温曲线 (**合理**) (2分)

**问题二 (15分)**

**模型分 5分**

根据制程界限, 建立使传送带过炉速度最大的优化模型 (3-5分)

**算法分 5分**

有优化模型求解计算得流程图, 有对应程序代码, 写清楚计算步骤 (3-5分)

**结果分 5分**

最大传送带过炉速度 80 (+5) (3-5分)

**问题三 (20分)**

**模型分 5分**

给出超过 217°C 到峰值温度所覆盖的面积表达式, 并建立使该面积最小的优化模型, (3-5分)

**算法分 10分**

给出求解方法的流程图或代码或具体的求解方法, 以及算法中所涉及到的参数设置。 (3-10分)

**结果分 5分**

最大传送带过炉速度 面积 500-800 (3-5分)

**问题四 (15分)**

**模型分 5分**

给出超过 217°C 的炉温曲线对称性的数学描述, 建立考虑对称性及面积最小的多目标优化模型 (双目标优化模型) (3-5分)

**算法分 5分**

有优化模型求解计算得流程图, 有对应程序, 写清楚计算步骤 (3-5分)

**结果分 5分**

给出各温区设定温度和传送带过炉速度, 以及相应的指标值。 (3-5分)

# 2020C题（统计分析类） 中小微企业的信贷决策

在实际中，由于中小微企业规模相对较小，也缺少抵押资产，因此银行通常是依据信贷政策、企业的交易票据信息和上下游企业的影响力，向实力强、供求关系稳定的企业提供贷款，并可以对信誉高、信贷风险小的企业给予利率优惠。银行首先根据中小微企业的实力、信誉对其信贷风险做出评估，然后依据信贷风险等因素来确定是否放贷及贷款额度、利率和期限等信贷策略。

某银行对确定要放贷企业的贷款额度为10~100万元；年利率为4%~15%；贷款期限为1年。附件1~3分别给出了123家有信贷记录企业的相关数据、302家无信贷记录企业的相关数据和贷款利率与客户流失率关系的2019年统计数据。该银行请你们团队根据实际和附件中的数据信息，通过建立数学模型研究对中小微企业的信贷策略，主要解决下列问题：

(1) 对附件1中123家企业的信贷风险进行量化分析，给出该银行在年度信贷总额固定时对这些企业的信贷策略。

(2) 在问题1的基础上，对附件2中302家企业的信贷风险进行量化分析，并给出该银行在年度信贷总额为1亿元时对这些企业的信贷策略。

(3) 企业的生产经营和经济效益可能会受到一些突发因素影响，而且突发因素往往对不同行业、不同类别的企业会有不同的影响。综合考虑附件2中各企业的信贷风险和可能的突发因素（例如：新冠病毒疫情）对各企业的影响，给出该银行在年度信贷总额为1亿元时的信贷调整策略。

## C题 中小微企业的信贷决策

评分项目	评分内容	说明	分数
论文摘要	问题1, 问题2, 问题3 建模方法, 计算方法, 问题答案	具体写出	10分
问题1 (30分)	对数据进行预处理并分析, 得到不同企业数据的差异性并可视化。	数据分析	5分
	依据附件1中相关数据, 提出量化企业信贷风险的指标, 包括综合实力和信誉, 综合分析给出结论。	计算方法	10分
	将各企业进行风险分类, 建立基于贷款利率与客户的流失率的信贷决策模型, 给出最优的贷款额度、利率及分配占额策略。	模型 算法	10分
	将每个企业的信贷策略结果可视化, 分析合理, 特别是对有违约纪录的企业不予放贷。	策略分析	5分
问题2 (35分)	利用附件1中所有企业的交易数据做机器学习, 预测附件2中企业的违约概率, 从而得到企业的信誉评级。	算法描述 清楚	15分
	参照问题1的方法, 量化企业的经营实力, 对企业信誉进行评级。以此建立信贷风险量化模型, 并给出量化分析结果。	模型描述 清楚	10分
	在风险量化模型的基础上, 建立信贷决策模型, 给出银行总额为1亿元时的信贷策略。	策略合理	10分
问题3 (15分)	考虑新冠疫情或其他突发因素对不同行业企业的影响。	模型描述 清楚	10分
	在上述突发因素影响下, 重新评估企业的信贷风险, 调整银行信贷决策, 分析影响及效果。	策略合理	5分
论文写作	语言, 排版, 完整等	印象	5分
	模型检验, 模型稳定, 参数影响	加分说明	
	模型特色	问题3建立了风险异常检测模型	5分

- 摘要、论文写作（印象）：15分
- 数据分析、指标的确定
- 信贷决策、信贷风险量化模型及算法
- 结果（合理、分析）
- 论文特色加分：5分

# 数学建模论文评阅，重点关注以下几点

**形式上：**要素齐全、版面整齐美观；论文可读性强

**内容上：**表述规范、准确、到位；逻辑思路清晰；凸显基于问题驱动的数学建模、基于模型驱动的实现；回答问题、重点突出

- 论文摘要、论文的写作（文章的可读性等）
- 数据收集、分析和处理、评价指标确定（权威性、合理性）
- 问题的假设、基于问题驱动的建模过程（合理性、创新）
- 基于模型和应用场景的算法设计（针对性、创新）
- 结果分析、模型和算法的评价（完整性）
- 文献引用（规范性）



2019A题 高压油管的压力控制

高压油管的压力控制

摘要

高压油管压力控制是保证发电机安全运行的必要手段。本文依据质量守恒定律，建立了**燃油密度推导方程**，通过**差分法**迭代出任意时刻高压油管燃油密度，提供燃油压力稳定时的最优供油策略。

针对问题一，首先建立燃油密度与压强的积分方程，通过积分离散化求解出燃油密度与压强的关系式。为维持高压油管燃油压力不变，即周期内高压油管燃油质量进出相同。先根据液体流速公式计算进出高压油管燃油体积，将高压油泵、高压油管初始状态带入质量守恒方程可求得燃油密度推导方程。建立**积分模型**，再对积分方程进行**离散化**，用 *MATLAB* **遍历搜索**计算出单向阀开启时间为 **0.289 ms**时，高压油管压力稳定在 100 MPa。稳压 150 MPa 时单向阀开启时间如下表：

稳压 150 MPa 用时	2 s	5 s	10 s
单向阀开启时间	14.39 ms	3.7 ms	1.76 ms

针对问题二，首先结合喷油嘴针阀运动规律与液体流速公式，计算出喷油速率分布。再由凸轮机极角极径尺寸，确定凸轮机角速度与高压油泵燃油密度变化关系。依据质量守恒定理建立高压油管内部燃油密度递推方程式，建立**最优角速度模型**。通过对积分方程进行离散化，进行求解得到凸轮的角度速度。用 *MATLAB* **遍历搜索**计算出稳压 100 MPa 状态下凸轮机角速度为 **0.316 rad/ms**。

针对问题三，首先对比两喷油嘴同时工作与交替工作（即工作间隔 50ms），高压油管内部燃油压力变化图。建立**最优角速度模型**，求解得到凸轮的角度速度，两喷油嘴交替工作下，输油速率更均匀，燃油压力相较 100 MPa 残差和更小，即控油效果稳定性更优。

经分析得出压力残差的本质是燃油进出速率分布不均匀的结果。当高压油管内部燃油压力大于 100 MPa，启用单向减压阀；反之，关闭。此时，由于高压油管进油与单向减压阀排油的时间分布基本重叠，通过调节凸轮机转速，可使燃油进出高压油管速率更为接近。结果如下：

	双喷嘴同时工作	双喷嘴间隔 50 ms 工作	双喷嘴和减压阀同时工作
残差和	202900 MPa	99100 MPa	10200 MPa
幅值	0.41 MPa	0.22 MPa	0.2 MPa

关键词： 质量守恒定理    积分方程    离散法    迭代残差和

基于动力学模型和混沌量子遗传算法的同心鼓策略优化

摘要

本文根据机理分析建立**动力学模型**，运用**实数编码混沌量子遗传算法**（RCQGA）求解多目标优化，给出不同优化目标的“同心协力”项目策略方案。

**针对问题一**，将整个颠球受力过程分成**两个阶段**进行分析求解，以每个人单次颠球竖直做功最小为目标求解最优协作策略。首先分析鼓面平行条件和绳子状态对鼓面稳定性的影响，得出当所有人员**等间距站立**，任意时刻绳与水平面夹角、所施加力相等时能够使鼓面保持平行；然后分鼓受绳子施力发生运动和发生碰撞后使球上升两个阶段进行受力分析，以竖直做功最小为目标建立理想状态下**最佳协作模型**，运用混沌量子遗传算法进行求解，**发力方向**沿绳与水平面的夹角为 **3.62°**，**发力时刻**为 **0.65+0.48(k-2)s**，**力度** 76.58N，该策略下**颠球高度**为 42cm。

**针对问题二**，将六个自由度的空间转动降维到单自由度旋转情况，选取转角为广义坐标，得到**拉格朗日方程**。将鼓边缘力的作用效果等效成两个作用力作用在局部几何中心的效果，通过**虚功原理**给出广义力和转动角度的关系。根据 9 种作用力的施力情况，将其分为同时不同力、同力不同时和既不同力也不同时**三类**，给出每类偏转角的求解方法。逐一代入相关参数求解，求得 9 种施力情况下 0.1 s 后的鼓面偏转角分别为：**0.31°、0.60°、0.25°、0.71°、1.56°、0.66°、0.94°、0.88°、0.88°**。

**针对问题三**，在问题二的基础上进一步分析实际情况由于施力者控力不当对鼓平衡造成的影响，通过增强发力过程中水平方向的作用效果来降低鼓面倾斜的风险。以施力过程**水平做功之和最大**为优化目标，以**动力学关系**为约束条件，建立单目标优化模型，用 RCQGA 算法求解出模型调整后，**发力方向**沿绳与水平面的夹角为 **5.04°**，**发力时刻**为 **0.8+0.6(k-2)s**，**力度** 75.91N，该策略下**颠球高度**为 43.43cm。

**针对问题四**，将调整阶段分为**两个阶段**。第一阶段目标是将球的运动方向调整到竖直方向，第二阶段目标是将倾斜的鼓面调整到水平。在问题二建立的微分方程模型基础上，反解出两个等效力和作用点位置的关系。再根据球倾斜投影方向将等效力分配给各个施力者，得到两个阶段 10 个作用力的大小和持续时间，并分析得到现实情形此策略的实施效果。

本文的优点在于：一，深入分析多力系统的作用机理，通过建立**拉格朗日动力学方程**，降低矢量分析的求解难度；二，引入等效力，简化了多力的分析过程。  
关键词： 控制模型    动力学分析    多目标优化    混沌量子遗传算法

### 5.3.3 约束条件

#### 1、拉绳人员之间间隔约束

由于两拉绳人员之间的最小距离不小于 60cm，即

$$(r + l \cos \varphi) \sin \frac{\pi}{n} > 0.6 / 2, \quad (5-11)$$

其中， $r$  为鼓面半径 (0.2m)， $l$  为绳子长度， $n$  为团队人员人数， $2\varphi$  为两队员与鼓圆心连线的夹角。

#### 2、发力间隔约束

鼓与球碰撞分离后至下一次发力之间的时间间隔  $\Delta t_1$  要大于球与鼓碰撞分离后上升至最高点后下降到下一次需要发力之间的时间间隔  $\Delta t_2$ ，即

$$\Delta t_1 > \Delta t_2. \quad (5-12)$$

#### 3、球和鼓走过距离约束

项目开始时，球从鼓面中心上方 40cm 处竖直落下，球与鼓相碰时其分别走过距离  $s_1$ 、 $s_2$  之和要等于 40cm，即

$$s_1 + s_2 = 0.4m. \quad (5-13)$$

#### 4、颠球高度约束

由于球被颠起高度应开鼓面 40cm 以上，则

$$H < 0.4m. \quad (5-14)$$

### 5.3.4 单次颠球每个人最小做功目标

实际中一次颠球过程中团队中每个人竖直做功尽可能少即耗散能量少才能使颠球次数尽可能多，当保持鼓面平行时，每个人一次颠球过程中所做功  $W$  为

$$W = F \cdot \sin \theta \cdot x. \quad (5-15)$$

### 5.3.5 最佳协作策略模型

以一次颠球过程中团队中每个人做功  $W$  尽可能少，结合各约束条件，给出最佳配合策略，并给出该策略下的颠球高度  $H$ ：

$$\min W \quad (5-16)$$

## 优化模型建立过程，如：

分析问题，并  
引入决策变量



目标函数的分析



约束条件的分析



$$\begin{aligned} \max \quad & I \left( \sum_{i=1}^{24} S_i \eta_i x_i \right) \eta C D Y - \sum_{i=1}^{24} c_i W_i x_i \\ & - \text{逆变器费用,} \\ \text{s.t.} \quad & \sum_{i=1}^{24} W_i x_i \geq 8000, \\ & \sum_{i=1}^{24} S_i x_i \leq S, \\ & x_i \geq 0 \text{ 并取整数,} \end{aligned}$$

增强论文的可读性（至关重要）

## 2.2 数模论文中常见的问题

**可读性：**文字表述不准确、不规范（口语化）；  
论文排版的不规范、不美观；建模过程与模型呈现（不明确、建模思路不清晰、模型呈现突然）；  
数据结果的呈现（不规范、重点不突出、杂乱）

**逻辑性：**在模型建立（推导）、算法设计（使用）等方面，“流水账”式的直接呈现多（What），  
“讲道理（Why）”少



表 1.2 各面的平均辐射强度 (单位: 瓦 / (天 · 米<sup>2</sup>))

门限	顶面	东面	南面	西面	北面
80	4.2776	1.4271	2.7268	2.1810	0.3617
30	4.3364	1.5903	2.8656	2.3795	0.6651

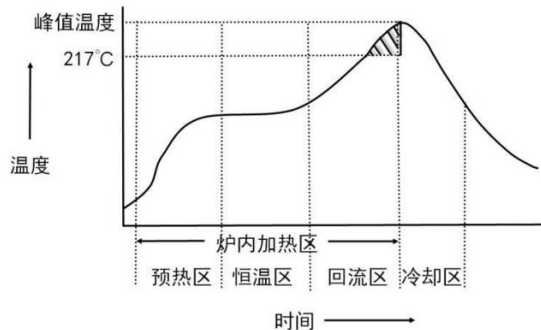


图 2 炉温曲线示意图

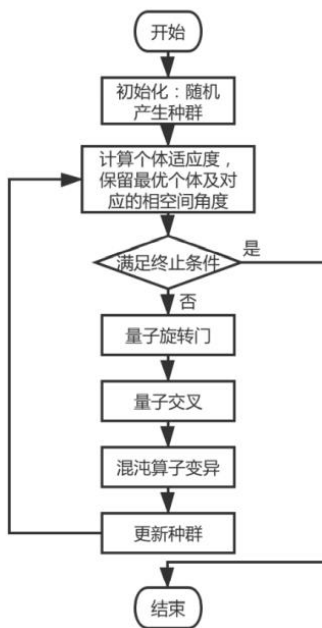
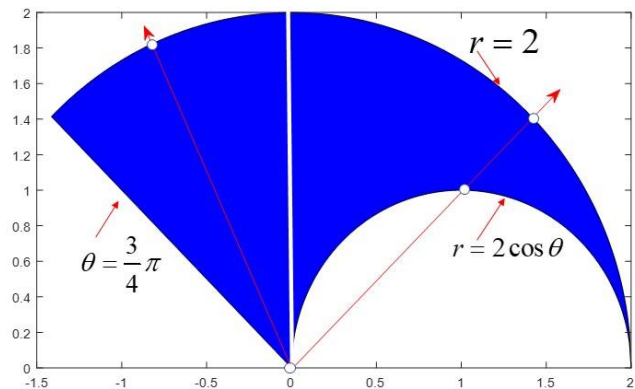


图 5.3 混沌量子遗传算法流程图



$$\int_{-\sqrt{2}}^0 dx \int_{-x}^{\sqrt{4-x^2}} (x^2 + y^2) dy + \int_0^2 dx \int_{\sqrt{2x-x^2}}^{\sqrt{4-x^2}} (x^2 + y^2) dy$$

What? How?



### 三、符号说明

符号	意义	单位
$v_A'$	$t$ 时刻燃油经过高压油泵从 $A$ 处进入高压油管的体积	$mm^3$
$v_B'$	$t$ 时刻燃油经过高压油管从喷孔 $B$ 喷出油的体积	$mm^3$
$v_{\text{管}}$	高压油管的柱形容积	$mm^3$
$p_{\text{泵}}'$	$t$ 时刻高压油泵的压力	$MPa$
$p_{\text{管}}'$	$t$ 时刻高压油管的压力	$MPa$
$p_{\text{低}}$	低压燃油的压力	$MPa$
$c_A(t)$	燃油经过高压油泵从 $A$ 处进入高压油管喷油速率	$mm^3/ms$
$c_B(t)$	燃油经过高压油管从喷孔 $B$ 喷油速率	$mm^3/ms$
$c_D(t)$	单向减压阀 $D$ 喷油速率	$mm^3/ms$
$t_0$	单向阀单位周期内开启的时长	$ms$
$\rho_{\text{管}}'$	$t$ 时刻高压油管燃油的密度	$mg/mm^3$
$\rho_{\text{泵}}'$	$t$ 时刻高压油泵燃油的密度	$mg/mm^3$

(注：未列出符号及重复符号以出现处为准)

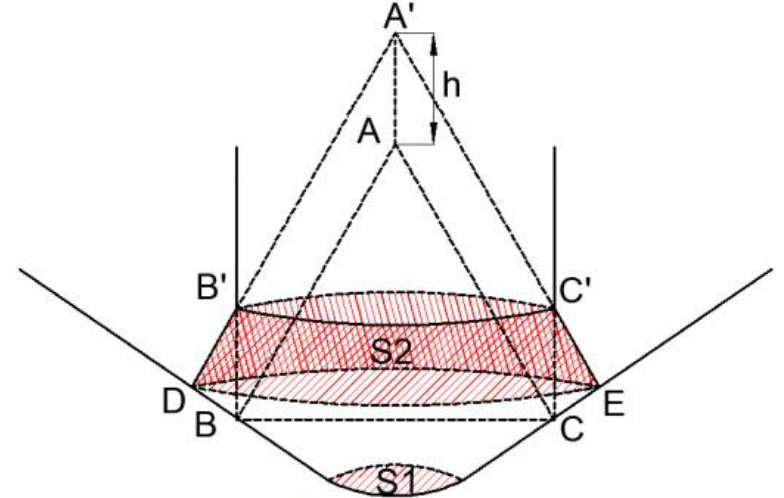


图 4：喷油器喷嘴孔示意图

#### (4) 动量守恒和能量守恒

对球和鼓的碰撞时刻  $t_1$  进行动量守恒分析，碰撞时间忽略不计，下标 1 为球，下标 2 为鼓， $v_1'$  和  $v_2'$  分别为球和鼓碰撞后的速度，方程如下：

$$m_1 v_1(t_1) + m_2 v_2(t_1) = m_1 v_1'(t_1) + m_2 v_2'(t_1)$$

由于碰撞时能量损失远小于碰撞时鼓对球做的功，故可建立能量守恒方程，如下：

$$\frac{1}{2} m_1 v_1^2(t_1) + \frac{1}{2} m_2 v_2^2(t_1) = \frac{1}{2} m_1 v_1'^2(t_1) + \frac{1}{2} m_2 v_2'^2(t_1)$$

上两式即构建碰撞前后球和鼓的 4 个速度之间的函数关系。

#### (5) 颠球高度

对球碰撞后到上升至最高点（高度大于 0.4m），列出动力学方程：

$$x_1' = \frac{v_1'^2}{2g} \quad (x_1' \geq 0.4)$$

通过上式，计算出碰撞后球得速度  $v_1'$  后，即可求得颠球高度  $x_1'$ 。

综合可得单目标优化模型为：

$$\begin{aligned} \min_x \quad & z = \frac{\sum_{i=1}^n \left| \frac{P_{\max}^i + P_{\min}^i}{2} - 100 \right|}{n} \\ \text{s.t.} \quad & \begin{cases} P_{\text{pipe}} = P \left( \rho_{\text{pipe}}(0) + \int_0^t d\rho_{\text{pipe}} \right) \\ d\rho_{\text{pipe}} = \frac{dm_{\text{pipe}}}{V_{\text{pipe}}} \\ dm_{\text{pipe}} = Q_{\text{in}}(t) \rho_{\text{pump}} dt - Q_{\text{out}}(t) \rho_{\text{pipe}}(t) dt \\ Q_{\text{in}}(t) = \begin{cases} CA \sqrt{\frac{2(160-P)}{\rho_{\text{pump}}}} & t \bmod (10+x) > 10 \\ 0 & t \bmod (10+x) \leq 10 \end{cases} \\ Q_{\text{out}}(t) = \begin{cases} 100t^* - 9760 & 97.6 \leq t^* \leq 97.8 \\ 20 & 97.8 < t^* \leq 99.8 \\ -100t^* + 10000 & 99.8 < t^* < 100 \end{cases} \quad (t^* = t \bmod 100) \end{cases} \end{cases} \quad (10)$$



标注不完整

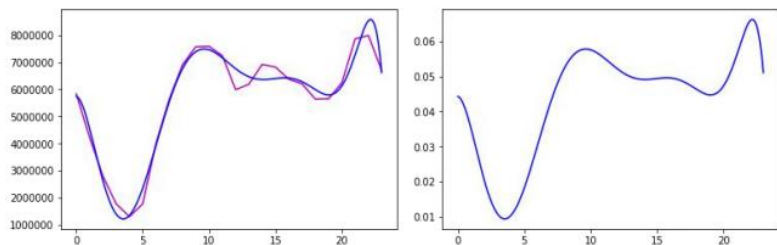


图 4 不同时间全市出租车里程及拟合归一化函数

$i$	0	1	2	3
$a_i$	0.044293843610886	-0.0007550934634	-0.011760684454	0.00366243206836
$i$	4	5	6	7
$a_i$	-0.00025139566161	-3.1041703933e-5	6.015863312e-6	-3.8790931755e-7
$i$	8	9		
$a_i$	1.146949531570e-8	-1.3111463356e-10	数据格式不规范	

表 7 不同时间出租车需求强度函数参数表

模糊不清

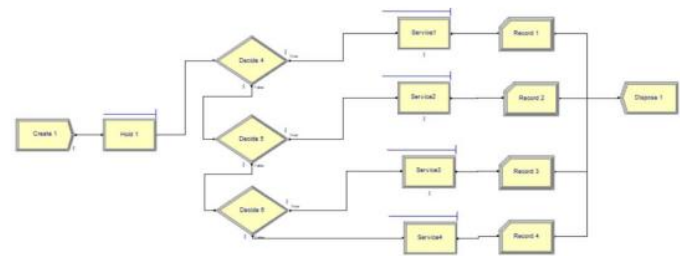
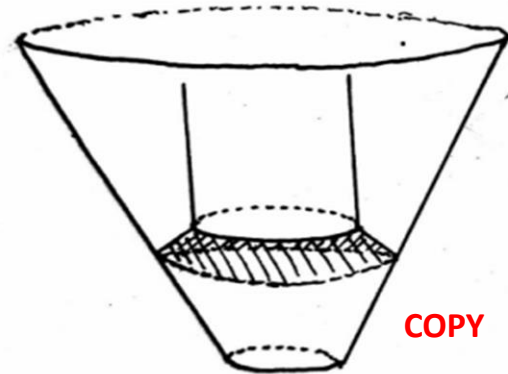


图 13 上车点“并行调度”仿真模型



(a) 燃油过针阀的截面示意图



2.2 问题 2 重述.....	7
2.3 问题 3 重述.....	7
2.4 问题 4 重述.....	7
2.5 问题 5 重述.....	7
2.6 问题 6 重述.....	7
3、问题分析 .....	8
3.1 问题 1 分析.....	8
3.2 问题 2 分析.....	8
3.3 问题 3 分析.....	8
3.4 问题 4 分析.....	8
3.5 问题 5 分析.....	8
3.6 问题 6 分析.....	9
4、模型构建 .....	9
4.1 混合高斯模型 (GMM) .....	9
4.2 RPCA 模型及 ALM 算法 .....	10
4.3 在线混合高斯低维矩阵分解(OMoGMF) .....	11
4.4 基于显著程度的视频关键帧提取模型.....	13
4.5 SIFT 特征匹配算法.....	15
5、问题求解 .....	16
5.1 问题 1 的求解.....	16
5.2 问题 2 的求解.....	19
5.3 问题 3 的求解.....	19
5.4 问题 4 的求解.....	21
5.5 问题 5 的求解.....	22
5.6 问题 6 的求解.....	26





1. 问题重述.....	4
1.1 问题背景.....	4
1.2 基本问题.....	4
2. 模型假设与符号说明.....	5
2.1 模型假设.....	5
2.2 符号说明.....	5
3. 问题一分析与求解.....	5
3.1 问题分析.....	5
3.2 问题求解.....	8
3.3 小结.....	9
4. 问题二分析与求解.....	9
4.1 问题分析.....	9
4.2 问题建模与求解.....	12
4.3 小结.....	13
5. 问题三分析与求解.....	14
5.1 问题分析.....	14
5.2 问题求解.....	14
5.3 小结.....	16

## 2.3 写好数模论文

- 充裕的论文写作时间【时间安排】
- 三名队员间的协作【论文、建模、算法&编程】，提高论文撰写效率和质量
- 模拟训练【建模、算法、论文写作：准备】、优秀数模论文范例学习
- 提升数模论文质量和竞赛成绩

### 三、建模案例：高精度参数的确定问题

- 问题的提出

设有一生态系统，其中含有两种生物，A 生物是捕食者和 B 生物是被捕食者，假设  $t$  时刻捕食者 A 的数目为  $x(t)$ ，被捕食者 B 数目为  $y(t)$ ，它们之间满足以下变化规律：

$$\begin{cases} \frac{dx(t)}{dt} = x(t)[\alpha_1 + \alpha_2 y(t)] \\ \frac{dy(t)}{dt} = y(t)[\alpha_3 + \alpha_4 x(t)] \end{cases}$$

初始条件为：  $\begin{cases} x(t_0) = \alpha_5 \\ y(t_0) = \alpha_6 \end{cases}$  , 其中  $\alpha_k (1 \leq k \leq 6)$  为模型的待定参数。

# 观测数据

通过对此生态系统的观测，可以得到相关的观测数据：  
在观测数据无误差的情况下

DATA1.txt

观测时刻 $t_j$	A 生物数目 $x(t_j)$	B 生物数目 $y(t_j)$
0	10.	60.
0.1	11.750840650304518	0.1374480266382216
0.2	3.4133367257849176	7.108705996120129
0.30000000000000004	20.80921881438798	0.4251595082899424
0.4	5.231982481945481	0.7182384931413827
0.5	26.925221604818102	26.706603701525214

在观测资料有误差（时间变量不含有误差）的情况下

DATA2.txt（情形一）

观测时刻 $t_j$	A 生物数目 $x(t_j)$	B 生物数目 $y(t_j)$
0	12.962285633035274	72.12301583344777
0.1	7.556010311240527	71.67657541335356
0.2	5.012389496575022	48.23918464064572
0.30000000000000004	4.115101160280639	27.706526063384022
0.4	4.157519511393944	15.771296803359892
$\vdots$	$\vdots$	$\vdots$

DATA3.txt（情形二）

观测时刻 $t_j$	A 生物数目 $x(t_j)$	B 生物数目 $y(t_j)$
0	12.82305785349266	73.40728069086651
0.01	12.599270344184822	74.60773967393804
0.02	11.714887378529845	74.71135309095207
0.03	11.120995506688704	76.14504374543644
0.09	8.562924024759667	73.11730640827214
0.1	7.455022419608922	72.49767585236788
0.11	6.946599338970699	70.05667765725144
$\vdots$	$\vdots$	$\vdots$

# 请利用提供的数据，解决如下问题：

$$\begin{cases} \frac{dx(t)}{dt} = x(t)[\alpha_1 + \alpha_2 y(t)] \\ \frac{dy(t)}{dt} = y(t)[\alpha_3 + \alpha_4 x(t)] \\ x(t_0) = \alpha_5 \\ y(t_0) = \alpha_6 \end{cases}$$

- 1) 在**观测数据无误差**的情况下，若已知  $\alpha_2 = \frac{1}{5}$ ，求其它 5 个参数  $\alpha_k$  ( $k=1,3,4,5,6$ )。有关数据见数据文件：DATA1.TXT
- 2) 在**观测数据无误差**的情况下，若  $\alpha_2$  也未知，问至少需要多少组观测数据，才能确定参数  $\alpha_k$  ( $1 \leq k \leq 6$ )？有关数据见数据文件：DATA1.TXT
- 3) 在**观测资料有误差**（时间变量不含有误差）的情况下，请分别利用观测数据 DATA2.TXT 和 DATA3.TXT，确定参数  $\alpha_k$  ( $1 \leq k \leq 6$ ) 在某种意义下的最优解，并与仿真结果比较，进而改进数学模型。

# 微分方程反问题

$$\begin{cases} \frac{dx(t)}{dt} = x(t) [\alpha_1 + \alpha_2 y(t)] \\ \frac{dy(t)}{dt} = y(t) [\alpha_3 + \alpha_4 x(t)] \\ x(t_0) = \alpha_5 \\ y(t_0) = \alpha_6 \end{cases}$$

微分方程解：

$$\begin{cases} \text{A生物的数量: } x(t) = x(t, \alpha_1, \alpha_2, \dots, \alpha_6) \\ \text{B生物的数量: } y(t) = y(t, \alpha_1, \alpha_2, \dots, \alpha_6) \end{cases}$$


$$\frac{dx}{dy} = \frac{x(\alpha_1 + \alpha_2 y)}{y(\alpha_3 + \alpha_4 x)}$$

$$\frac{\alpha_3 + \alpha_4 x}{x} dx = \frac{\alpha_1 + \alpha_2 y}{y} dy$$

$$\alpha_3 \ln x + \alpha_4 x - \alpha_1 \ln y - \alpha_2 y = C$$

相空间轨线方程  
(不显含时间 $t$ )

# 第一问 ( $\alpha_2$ 已知)

  $\alpha_3 \ln x + \alpha_4 x - \alpha_1 \ln y - \alpha_2 y = C$

$$\alpha_3 \ln x_i + \alpha_4 x_i - \alpha_1 \ln y_i - C = \alpha_2 y_i, \quad i = 0, 1, \dots, 5$$

$$\begin{bmatrix} \ln x_0 & x_0 & -\ln y_0 & -1 \\ \ln x_1 & x_1 & -\ln y_1 & -1 \\ \vdots & \vdots & \vdots & \vdots \\ \ln x_5 & x_5 & -\ln y_5 & -1 \end{bmatrix} \begin{bmatrix} \alpha_3 \\ \alpha_4 \\ \alpha_1 \\ C \end{bmatrix} = \begin{bmatrix} \alpha_2 y_0 \\ \alpha_2 y_1 \\ \vdots \\ \alpha_2 y_5 \end{bmatrix} \rightarrow \text{最小二乘解, 即:}$$

$$\min_{\alpha_1, \alpha_3, \alpha_4, C} \sum_{i=0}^5 \{ \alpha_3 \ln x_i + \alpha_4 x_i - \alpha_1 \ln y_i - C - \alpha_2 y_i \}^2$$

$\alpha_1$	$\alpha_3$	$\alpha_4$	$C$
-2.00001342132307	12.0004164760665	-1.00004224667344	13.8203026133603
$\alpha_5 = 10, \alpha_6 = 60$			



$$\alpha_3 \ln x + \alpha_4 x - \alpha_1 \ln y - \alpha_2 y = C$$



或

将初始条件  $\begin{cases} x(t_0) = \alpha_5 \\ y(t_0) = \alpha_6 \end{cases}$  代入, 即得

$$C = \alpha_3 \ln \alpha_5 + \alpha_4 \alpha_5 - \alpha_1 \ln \alpha_6 - \alpha_2 \alpha_6$$

$$\alpha_3 (\ln x_i - \ln \alpha_5) + \alpha_4 (x_i - \alpha_5) - \alpha_1 (\ln y_i - \ln \alpha_6) - \alpha_2 (y_i - \alpha_6),$$

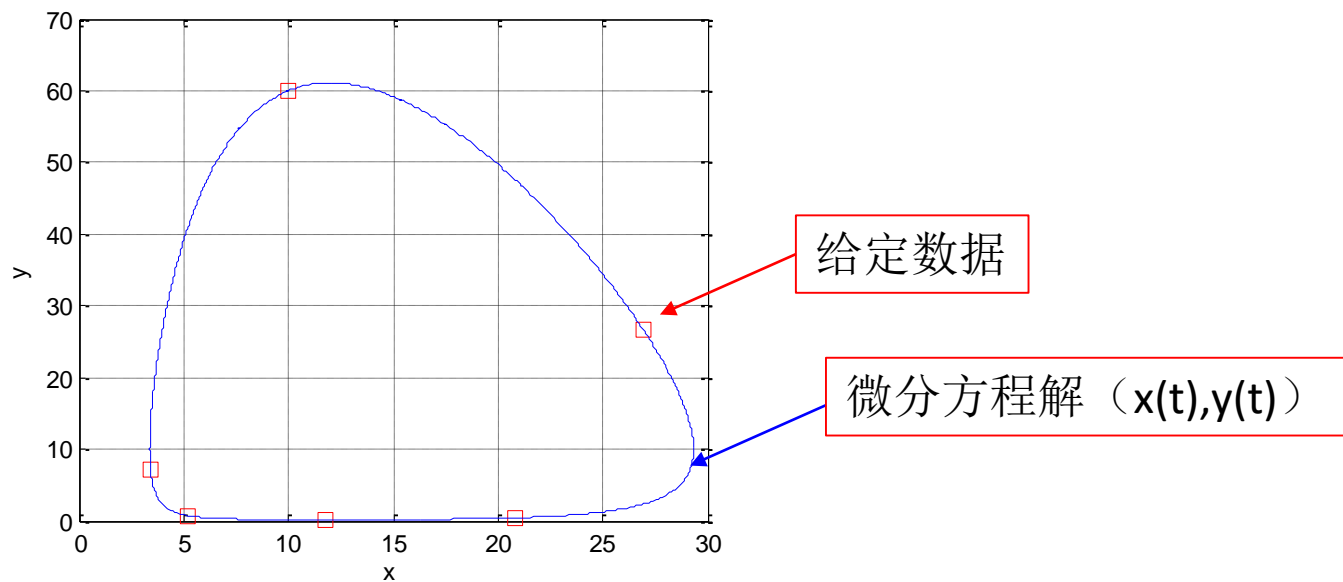
$$i = 0, 1, \dots, 5$$

最小二乘解, 即

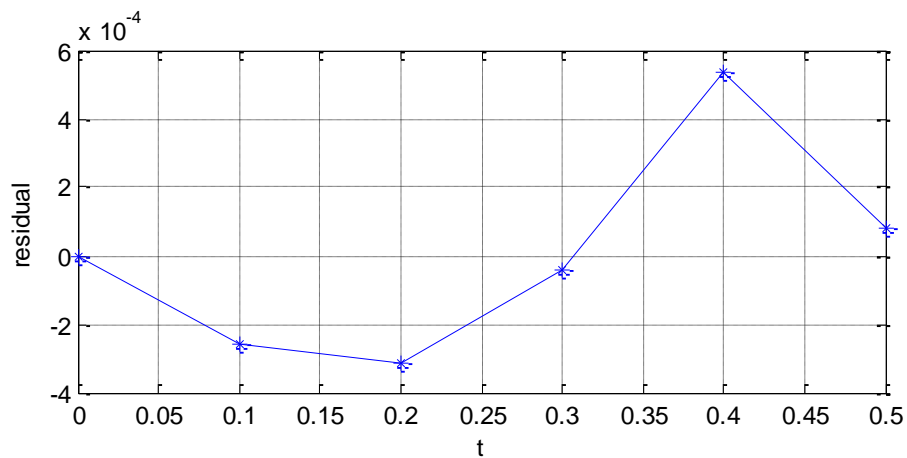
$$\min_{\alpha_1, \alpha_3, \alpha_4} \sum_{i=0}^5 \{ \alpha_3 (\ln x_i - \ln \alpha_5) + \alpha_4 (x_i - \alpha_5) - \alpha_1 (\ln y_i - \ln \alpha_6) - \alpha_2 (y_i - \alpha_6) \}^2$$

$\alpha_1$	$\alpha_3$	$\alpha_4$
-2.00001342132307	12.0004164760665	-1.00004224667344
$\alpha_5 = 10, \alpha_6 = 60$		

## 结果的检验



问题 1 相平面图



问题 1 误差

## 第二问 ( $\alpha_2$ 未知)

$$\alpha_3 \ln x + \alpha_4 x - \alpha_1 \ln y - \alpha_2 y = C \quad \frac{\alpha_3}{\alpha_2} \ln x + \frac{\alpha_4}{\alpha_2} x - \frac{\alpha_1}{\alpha_2} \ln y - \frac{C}{\alpha_2} = y$$

$$\text{或 } \frac{\alpha_3}{\alpha_2} (\ln x - \ln \alpha_5) + \frac{\alpha_4}{\alpha_2} (x - \alpha_5) - \frac{\alpha_1}{\alpha_2} (\ln y - \ln \alpha_6) = y - \alpha_6$$

在最小二乘意义下的  $\alpha_1/\alpha_2, \alpha_3/\alpha_2, \alpha_4/\alpha_2, C/\alpha_2$  的最优估计。

基于相轨线方程的线性最小二乘拟合求解结果

$\alpha_1/\alpha_2$	$\alpha_3/\alpha_2$	$\alpha_4/\alpha_2$	$C/\alpha_2$
-10.0000671066153	60.0020823803327	-5.00021123336719	69.1015130668015

在上述最优估计值已知的情况下,所谓  $\alpha_2$  的最优估计是指在此  $\alpha_2$  值下,

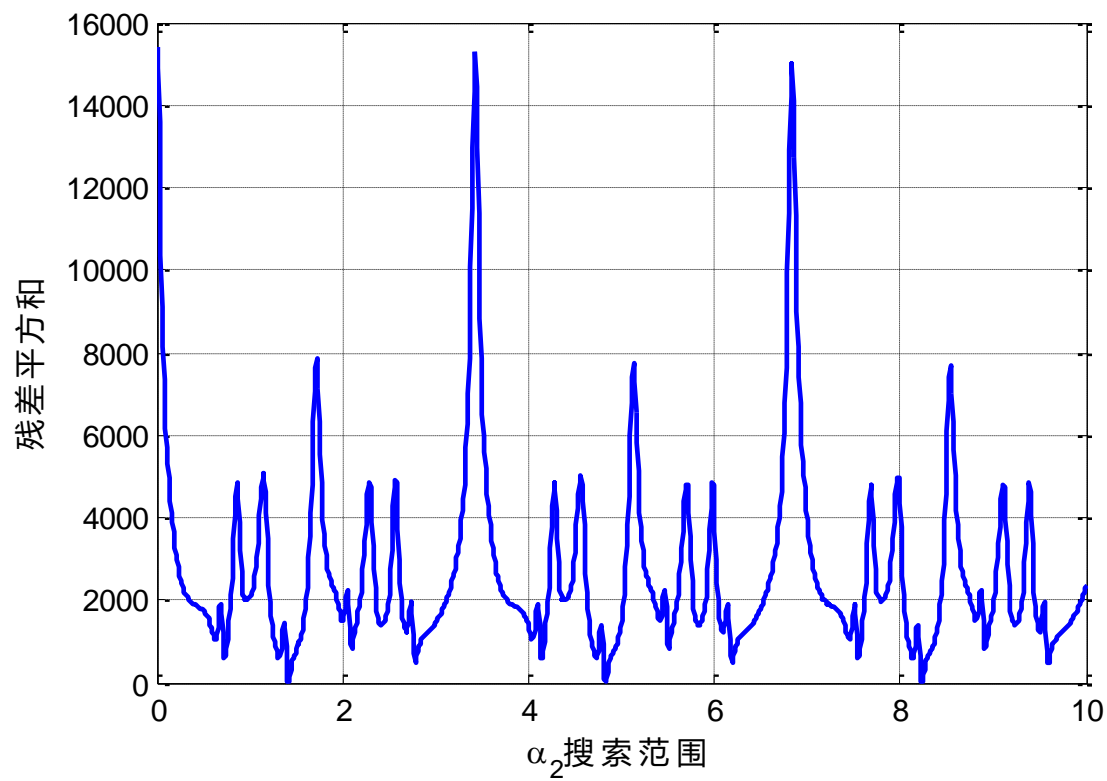
根据  $\alpha_1/\alpha_2, \alpha_3/\alpha_2, \alpha_4/\alpha_2$  的最优估计求出  $\alpha_1, \alpha_3, \alpha_4$ , 代入并求解常

微分方程, 求解结果与观测数据的误差最小

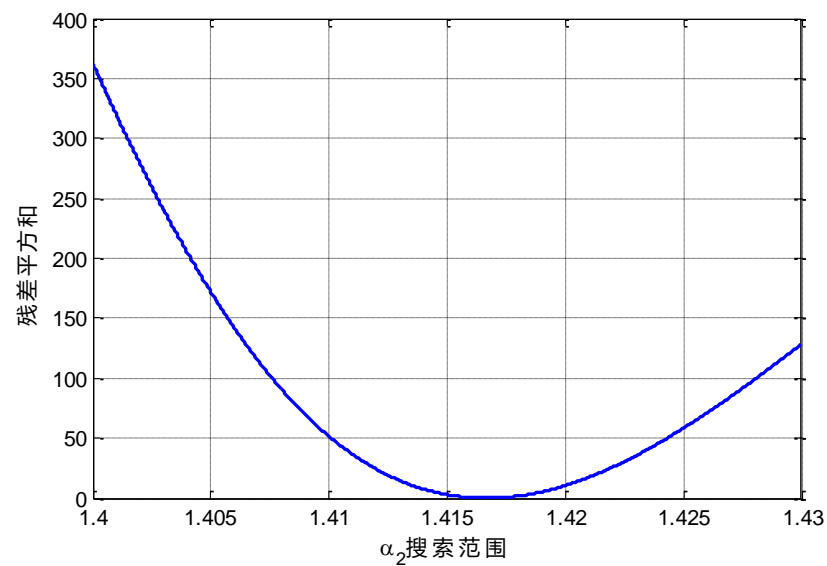
基于常微分方程数值求解的最优化模型：

$$\begin{aligned} \min_{\alpha_2} \quad & \sum_{i=0}^5 (x_i^p - x_i)^2 + \sum_{i=0}^5 (y_i^p - y_i)^2 \\ \text{s.t.} \quad & \begin{cases} \alpha_2 > 0 \\ \alpha_1/\alpha_2, \alpha_3/\alpha_2, \alpha_4/\alpha_2 \text{ (已知)} \\ \alpha_5, \alpha_6 \\ x_i^p, y_i^p \text{ 是给定系数 } \alpha_1, \alpha_2, \alpha_3, \alpha_4 \text{ 及初值 } \alpha_5, \alpha_6 \text{ 方程的解} \end{cases} \end{aligned}$$

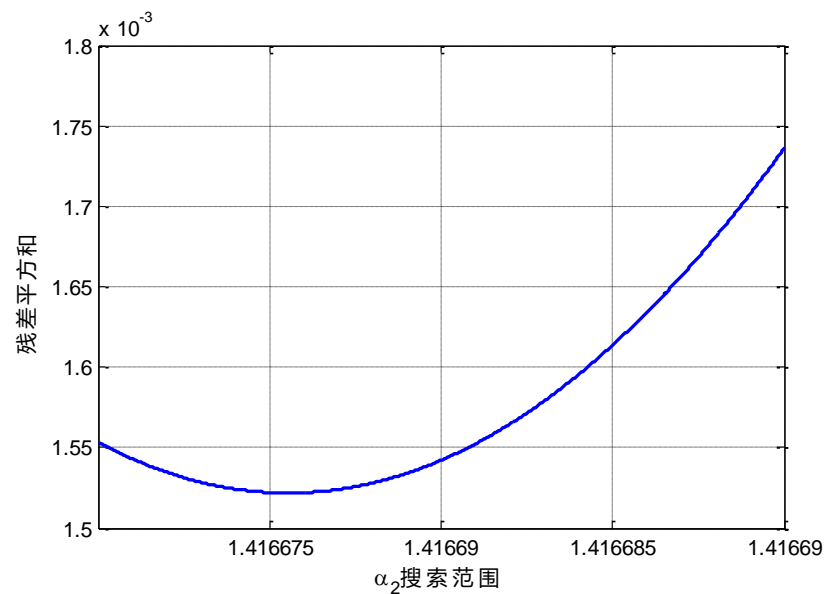
**如何求解：启发式搜索算法**



搜索范围[0,10]，搜索步长 0.01 对  $\alpha_2$  搜索的搜索结果



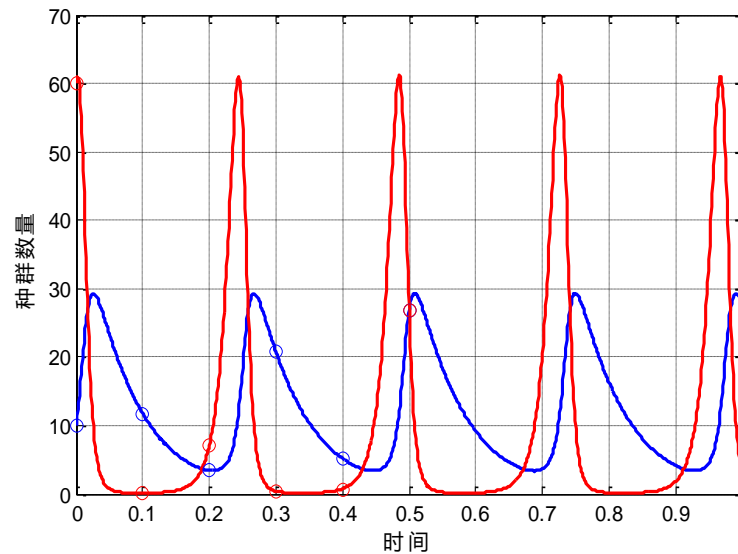
搜索范围[1.40,1.43]，搜索步长0.00001对  $\alpha_2$  搜索的搜索结果



搜索范围[1.41667-1.41669]，搜索步长0.00000001对  $\alpha_2$  搜索的搜索结果

DATA1.txt

观测时刻 $t_j$	A 生物数目 $x(t_j)$	B 生物数目 $y(t_j)$
0	10.	60.
0.1	11.750840650304518	0.1374480266382216
0.2	3.4133367257849176	7.108705996120129
0.30000000000000004	20.80921881438798	0.4251595082899424
0.4	5.231982481945481	0.7182384931413827
0.5	26.925221604818102	26.706603701525214



A生物 $x(t)$

B生物 $y(t)$

最优解对应常微分方程数值解与观测数据对照图

基于相轨线方程的线性最小二乘拟合至少需要 4 组数据，而最优化模型所需数据少于 4 组，因此，模型 2 至少需要 4 组数据

# 第三问（观测数据有误差）

在观测资料有误差（时间变量不含有误差）的情况下

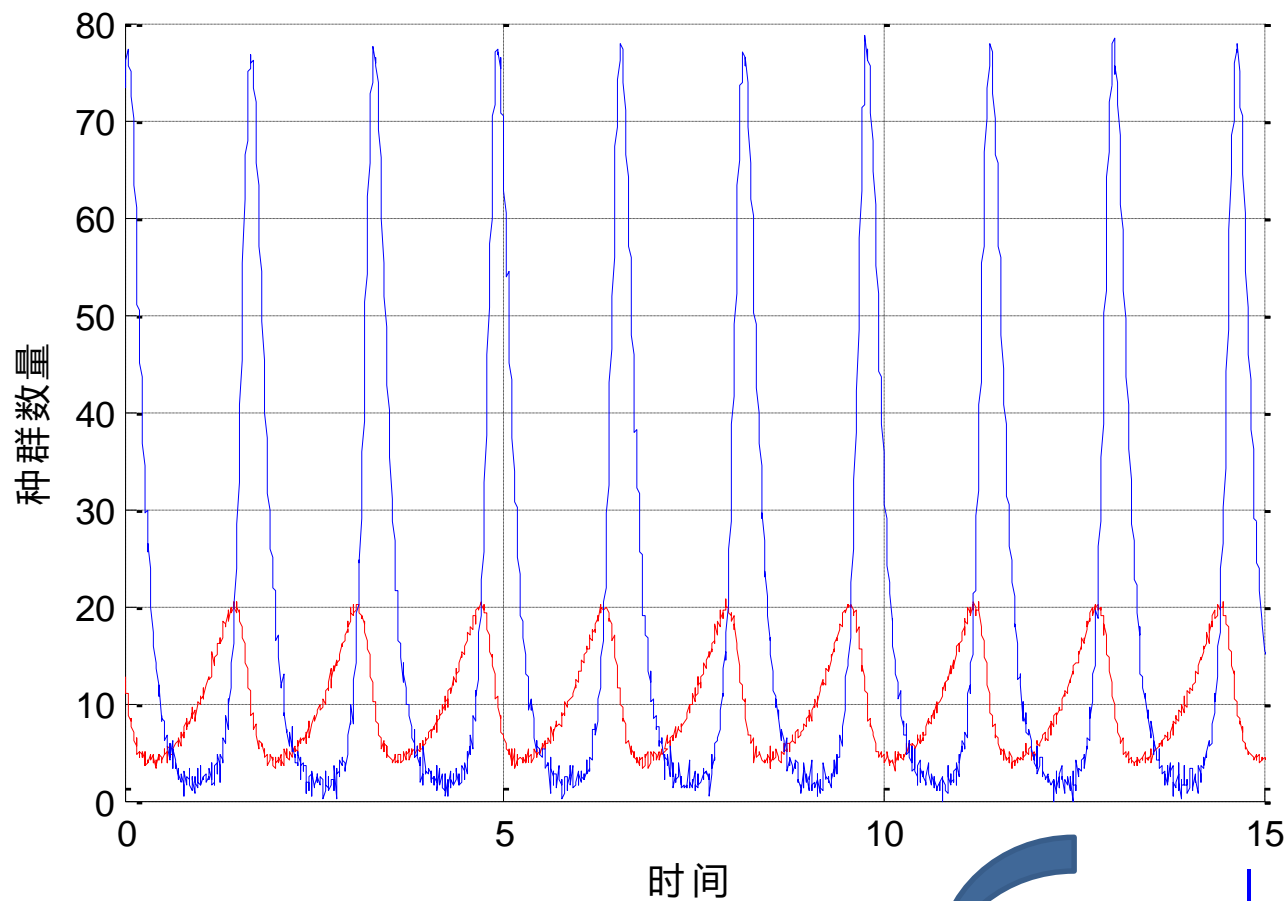
DATA2.txt（情形一）

观测时刻 $t_j$	A 生物数目 $x(t_j)$	B 生物数目 $y(t_j)$
0	12.962285633035274	72.12301583344777
0.1	7.556010311240527	71.67657541335356
0.2	5.012389496575022	48.23918464064572
0.30000000000000004	4.115101160280639	27.706526063384022
0.4	4.157519511393944	15.771296803359892
$\vdots$	$\vdots$	$\vdots$

DATA3.txt（情形二）

观测时刻 $t_j$	A 生物数目 $x(t_j)$	B 生物数目 $y(t_j)$
0	12.82305785349266	73.40728069086651
0.01	12.599270344184822	74.60773967393804
0.02	11.714887378529845	74.71135309095207
0.03	11.120995506688704	76.14504374543644
0.09	8.562924024759667	73.11730640827214
0.1	7.455022419608922	72.49767585236788
0.11	6.946599338970699	70.05667765725144
$\vdots$	$\vdots$	$\vdots$

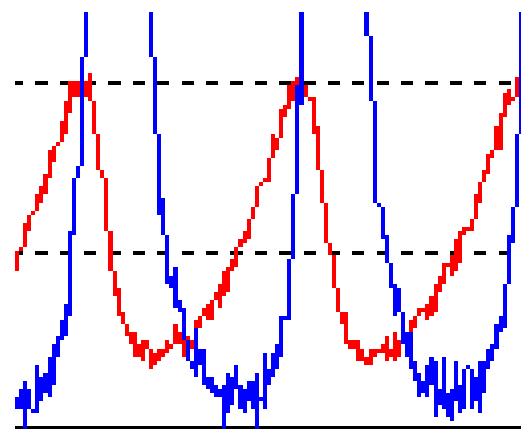




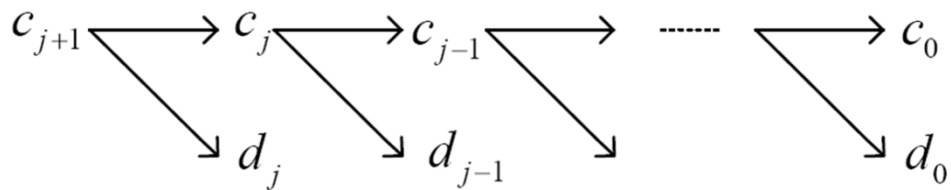
$A$ 生物 $x(t)$

$B$ 生物 $y(t)$

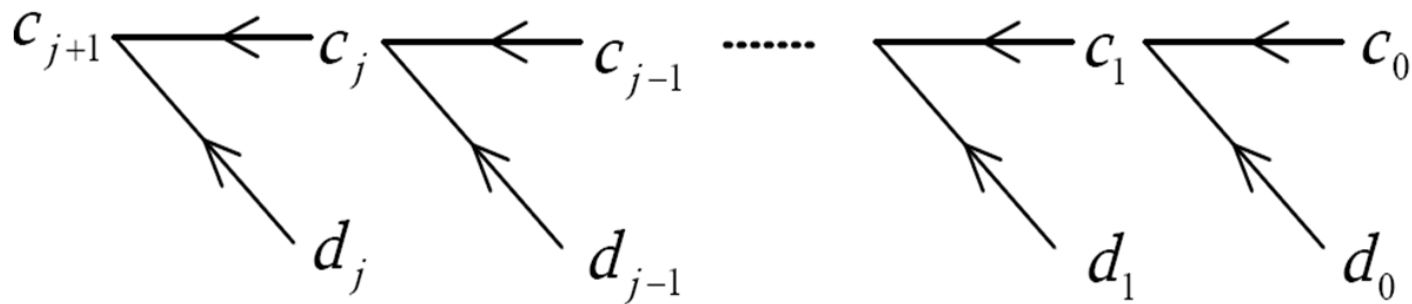
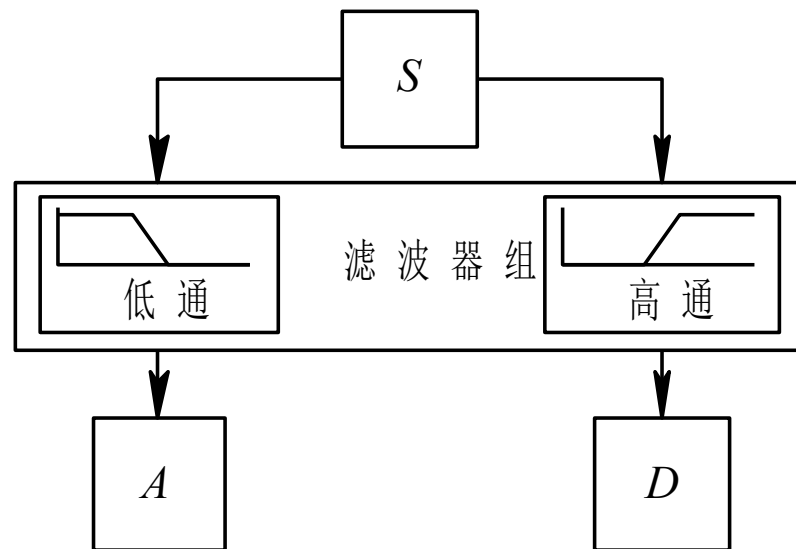
原始数据



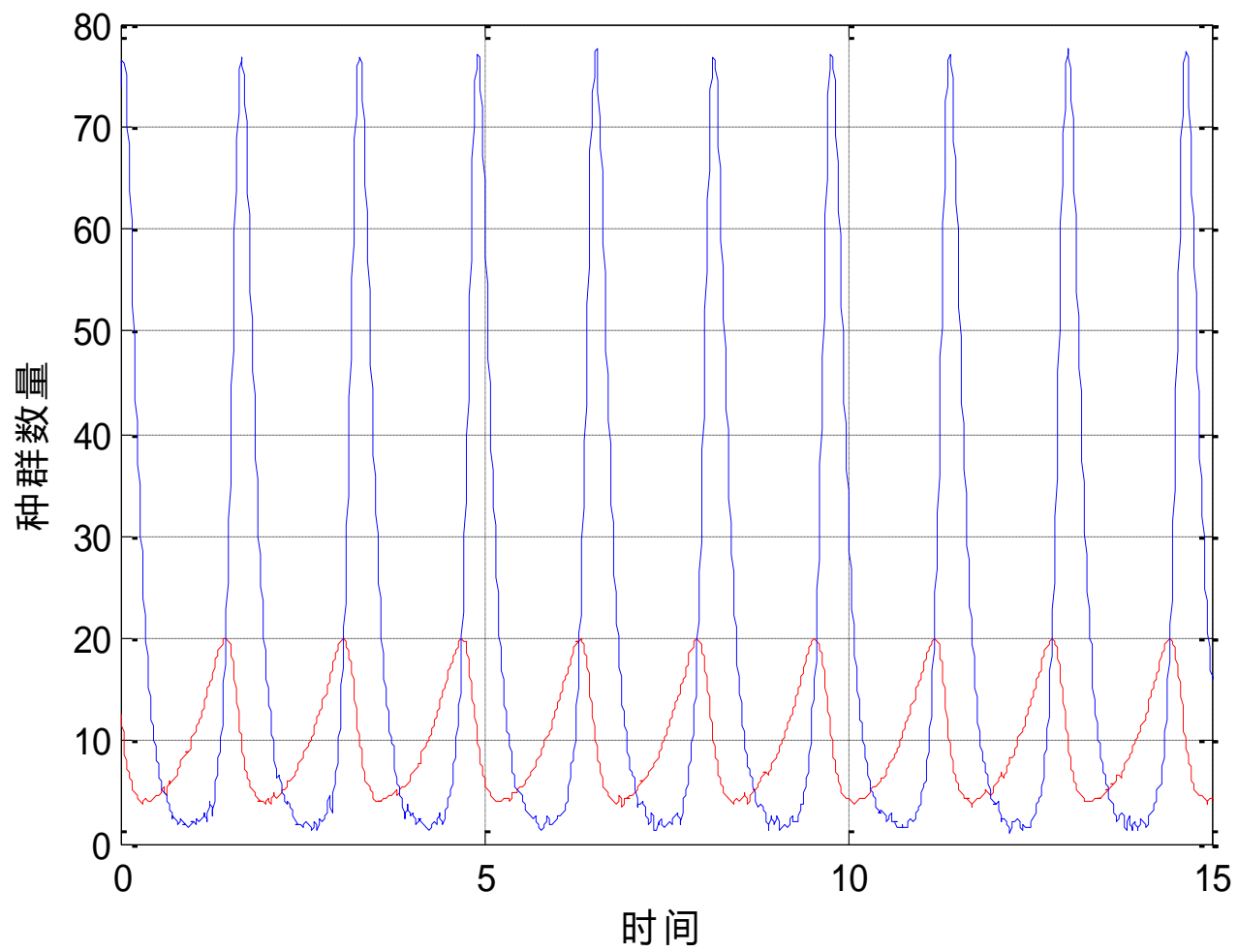
# 小波分析方法



Mallat 分解算法



Mallat 重构算法



去噪后的数据

考虑利用差商近似微商的方法建立关于  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$  的超定线性方程组（Euler 方法）

差分方程模型：

$$\frac{x(t_{i+1}) - x(t_i)}{t_{i+1} - t_i} \approx x(t_i)(\alpha_1 + \alpha_2 y(t_i))$$

$$\frac{y(t_{i+1}) - y(t_i)}{t_{i+1} - t_i} \approx y(t_i)(\alpha_3 + \alpha_4 x(t_i)) , \quad i = 0, 1, \dots, n-1$$

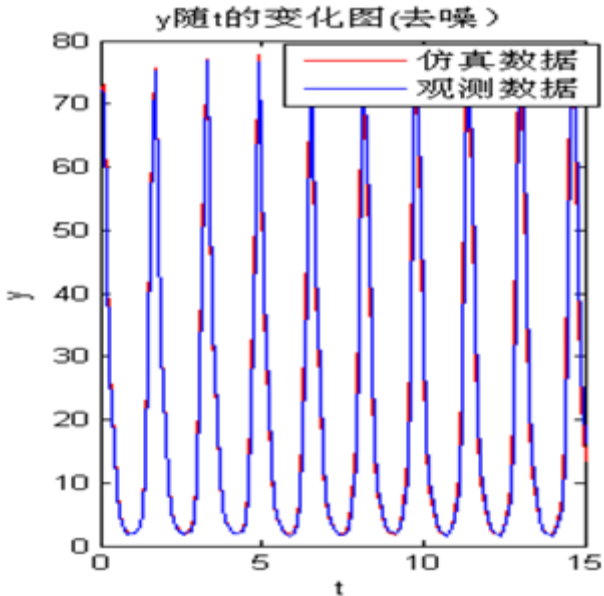
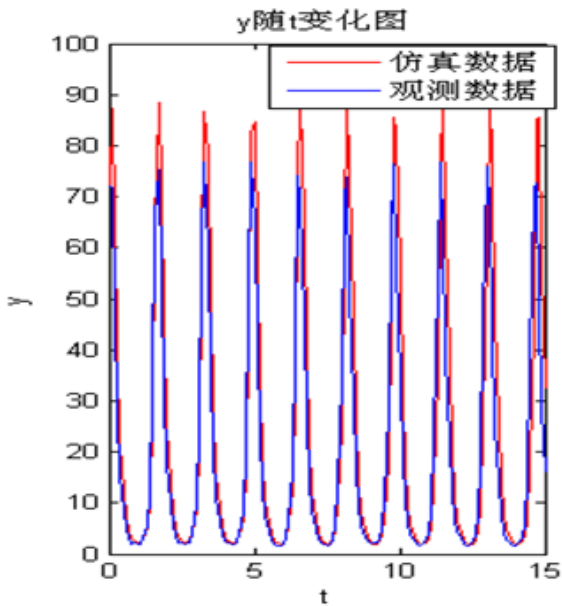
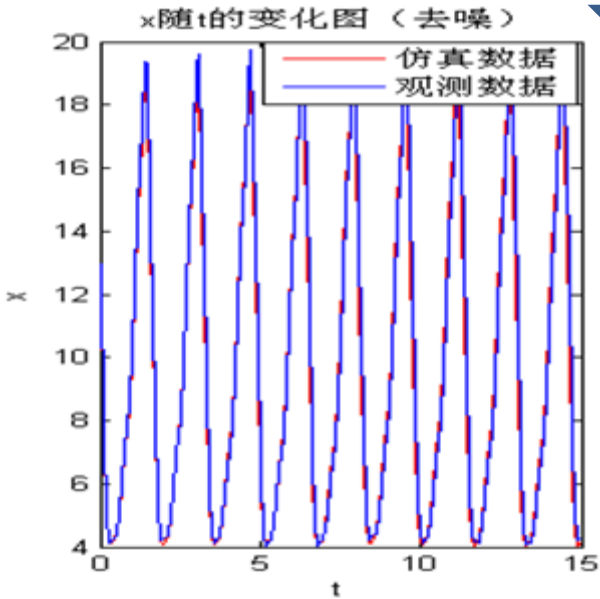
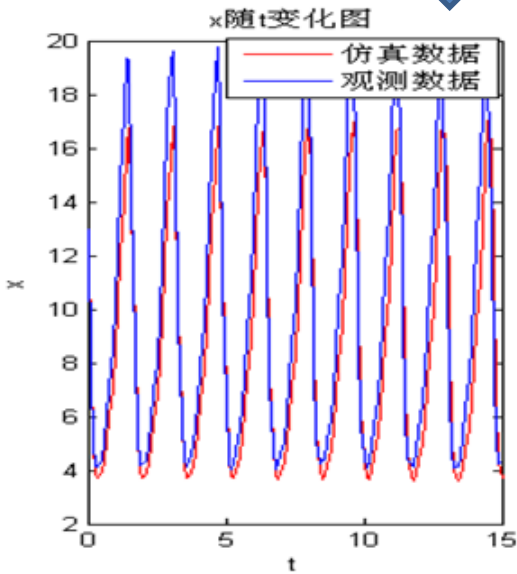
$\alpha_1, \alpha_2, \alpha_3, \alpha_4$  的求解结果如下表所示：

		$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	总体误差
Data2	去噪前	1.7182	-0.0881	-8.9734	0.8908	5.0719
	去噪后	2.0073	-0.0984	-9.9685	1.0123	1.9253
Data3	去噪前	-1.9887	0.1	9.9834	-0.9991	912.6333
	去噪后	2.0077	-0.0980	-9.9684	1.0124	76.1631

小波去噪前差分方法结果

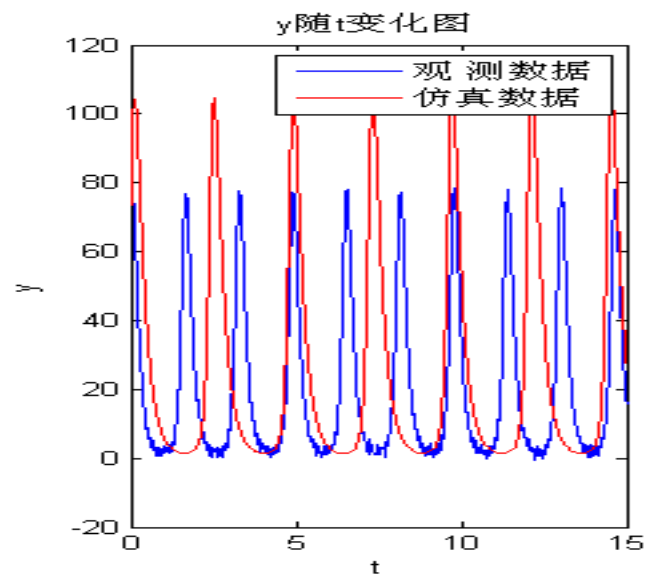
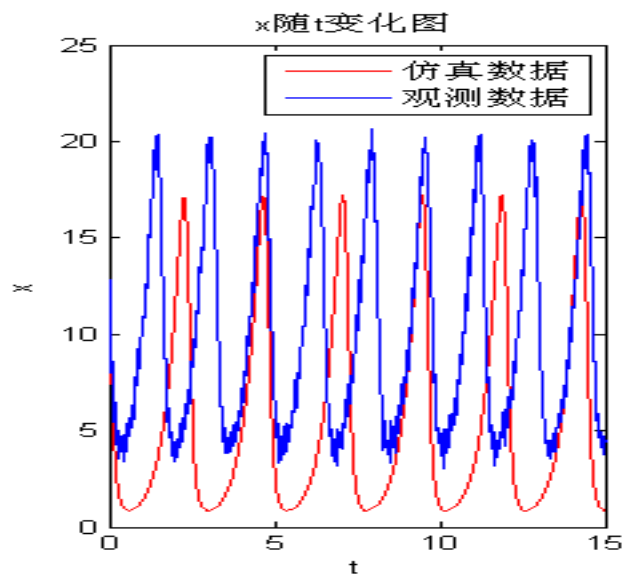


小波去噪后差分方法结果

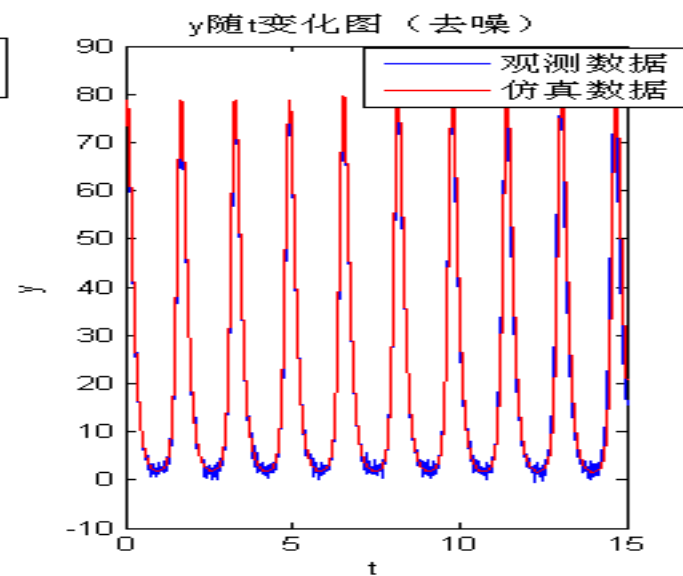
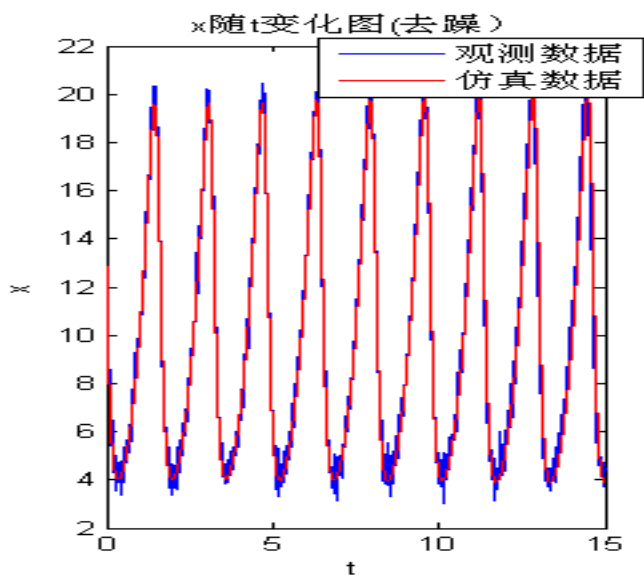


DATA3.TXT

小波去噪前差分法结果:



小波去噪后差分法结果:



# 微分方程反问题

## 最直接方法：利用求常微分方程数值解方法

微分方程初值问题，可以利用(给定  $\alpha_i, i=1,2,\dots,6$ )

matlab 中提供的函数 (ode23, ode45 等,  
即龙格库塔算法) 进行数值求解。

$$\begin{cases} \frac{dx(t)}{dt} = x(t)[\alpha_1 + \alpha_2 y(t)] \\ \frac{dy(t)}{dt} = y(t)[\alpha_3 + \alpha_4 x(t)] \\ x(t_0) = \alpha_5 \\ y(t_0) = \alpha_6 \end{cases}$$

$$\left\{ \begin{array}{l} x^p(t_i, \alpha_1, \alpha_2, \dots, \alpha_6) \\ y^p(t_i, \alpha_1, \alpha_2, \dots, \alpha_6) \end{array} \right\}, i=1,2,\dots,N, \text{ 利用给定数据 } \left\{ \begin{array}{l} x(t_i) \\ y(t_i) \end{array} \right\}$$

4) 利用最小二乘法

$$\begin{cases} \min_{\alpha_1, \alpha_2, \dots, \alpha_6} \sum_{i=1}^N \left\{ [x(t_i) - x^p(t_i, \alpha_1, \alpha_2, \dots, \alpha_6)]^2 + [y(t_i) - y^p(t_i, \alpha_1, \alpha_2, \dots, \alpha_6)]^2 \right\} \\ \text{s.t. : } \alpha_1, \alpha_2, \dots, \alpha_6 \text{ 可微函数.} \end{cases}$$

# 结语

## ①模型：基于问题驱动的数学建模

- **围绕问题展开&建模：**背景知识、存在的机理、类似问题的模型与求解方法（[查阅文献](#)）、收集相关的数据（[数据清洗](#)、[预处理](#)）、条件或要求转化为数学关系式【模型】：优化模型、统计分析模型、[层次分析模型](#)、微分方程模型、神经网络模型、灰色模型等等【模型的多样性】



## 基本建模（数学）方法的准备【正确掌握、正确使用、必要性】

- 最优化理论与算法、图论
- 统计方法（数据分析）
- 基本数值分析方法（插值、拟合、微分方程求解、数值积分等）
- 其它数学方法（层次分析法、预测、评价方法、神经网络、小波处理方法等）
- 相关背景专业知识的准备【参赛学生组队：专业交叉】

## ② 算法：基于模型驱动的设计

- 算法的选择：传统优化算法【贪婪算法、启发式算法：蜂群算法、遗传算法、模拟退火算法等】、统计分析、插值与拟合、常（偏）微分方程数值解、神经网络等等【基于模型、应用场景算法设计】

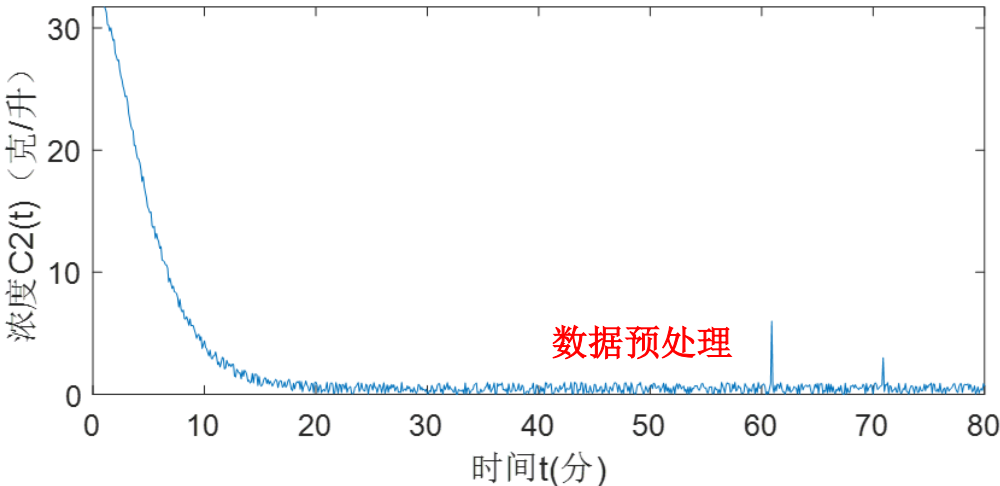
# 计算机语言编程及算法准备

- ◆ MATLAB \SPSS、R-软件\LINGO\Python\C语言（数学方法、算法的掌握与准备）
- ◆ 编程技巧、提高编程效率
- ◆ 常规算法实现（数据可视化、优化、统计分析、插值、拟合等）

# 数据预处理及数据使用方式

表 1. 测量数据。

时间 t	浓度C2(t)
1.0000000000000000	31.708621208344442
1.1000000000000000	31.602963404231442
1.2000000000000000	31.252706715623958
1.3000000000000000	30.278351435753837
1.4000000000000000	30.231365613577470
1.5000000000000000	29.741059947527337
1.6000000000000000	30.020363879163973
1.7000000000000000	29.431193685062329
1.8000000000000000	28.970637875040804
1.9000000000000000	29.052402214472156
.....	.....
79.4000000000000000	0.622332178602803
79.5000000000000000	0.103281768266669
79.5999999999999999	0.369867447279054
79.7000000000000000	0.501143611028728
79.7999999999999999	0.323648160632228
79.9000000000000000	0.895488435114924
80.0000000000000000	0.701990724632682



受噪音干扰的测量数据。

$$\alpha_3 \ln x + \alpha_4 x - \alpha_1 \ln y - \alpha_2 y = C$$

$$\frac{\alpha_3}{\alpha_2} \ln x + \frac{\alpha_4}{\alpha_2} x - \frac{\alpha_1}{\alpha_2} \ln y - \frac{C}{\alpha_2} = y$$

$$\frac{x(t_{i+1}) - x(t_i)}{t_{i+1} - t_i} \approx x(t_i)(\alpha_1 + \alpha_2 y(t_i))$$

使用方式（涉及数学、给定的条件、精度等考虑）

$$\frac{y(t_{i+1}) - y(t_i)}{t_{i+1} - t_i} \approx y(t_i)(\alpha_3 + \alpha_4 x(t_i)) , \quad i = 0, 1, \dots, n-1$$

### ③ 论文写作【获奖论文研读】

- 论文写作的基本要素：摘要、问题的“重述”、问题的分析、模型、算法、结果的分析与检验、推广、参考文献、主要程序
- 论文的可读性：版面及表述的规范、内容组织、结构层次、重点
- 突出基于问题驱动的建模思想（解决问题）
- 突出基于模型驱动的算法设计（可行性）
- 结果的**分析、检验、比较**等
- 论文的引用、查重
- 附录程序

**谢谢！**