

## 2013 高教社杯全国大学生数学建模竞赛

### 承 诺 书

我们仔细阅读了《全国大学生数学建模竞赛章程》和《全国大学生数学建模竞赛参赛规则》(以下简称为“竞赛章程和参赛规则”,可从全国大学生数学建模竞赛网站下载)。

我们完全明白,在竞赛开始后参赛队员不能以任何方式(包括电话、电子邮件、网上咨询等)与队外的任何人(包括指导教师)研究、讨论与赛题有关的问题。

我们知道,抄袭别人的成果是违反竞赛章程和参赛规则的,如果引用别人的成果或其他公开的资料(包括网上查到的资料),必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺,严格遵守竞赛章程和参赛规则,以保证竞赛的公正、公平性。如有违反竞赛章程和参赛规则的行为,我们将受到严肃处理。

我们授权全国大学生数学建模竞赛组委会,可将我们的论文以任何形式进行公开展示(包括进行网上公示,在书籍、期刊和其他媒体进行正式或非正式发表等)。

我们参赛选择的题号是(从 A/B/C/D 中选择一项填写): \_\_\_\_\_

我们的参赛报名号为(如果赛区设置报名号的话): \_\_\_\_\_

所属学校(请填写完整的全名): \_\_\_\_\_

参赛队员(打印并签名): 1. \_\_\_\_\_

2. \_\_\_\_\_

3. \_\_\_\_\_

指导教师或指导教师组负责人(打印并签名): \_\_\_\_\_

(论文纸质版与电子版中的以上信息必须一致,只是电子版中无需签名。以上内容请仔细核对,提交后将不再允许做任何修改。如填写错误,论文可能被取消评奖资格。)

日期: \_\_\_\_\_年\_\_\_\_月\_\_\_\_日

---

赛区评阅编号(由赛区组委会评阅前进行编号): \_\_\_\_\_

2013 高教社杯全国大学生数学建模竞赛

编 号 专 用 页

赛区评阅编号（由赛区组委会评阅前进行编号）：

赛区评阅记录（可供赛区评阅时使用）：

|     |  |  |  |  |  |  |  |  |  |  |
|-----|--|--|--|--|--|--|--|--|--|--|
| 评阅人 |  |  |  |  |  |  |  |  |  |  |
| 评分  |  |  |  |  |  |  |  |  |  |  |
| 备注  |  |  |  |  |  |  |  |  |  |  |

全国统一编号（由赛区组委会送交全国前编号）：

全国评阅编号（由全国组委会评阅前进行编号）：

# 城市表层土壤重金属污染与土壤自净能力研究

## 摘 要

摘要部分如何写，我们就不作介绍了，详细的按暑假写作课上给大家的魔法摘要模板和 7 步骤去完成。对于你们团队的讨论，切记“客观用词”和“6 小时”原则。

本文为解决第一问刻画某城区表层土壤 8 种重金属污染的空间分布问题，分别从定性和定量两个角度入手，运用差值拟合与多项式拟合，分别给出了 8 种重金属浓度在二维、三维空间上的分布图和反映该城区各重金属整体污染浓度同海拔区间的关系函数（见附录二）。

通过建立基于内梅罗指数法的重金属污染程度评价模型，结合该城区的表层土壤的 8 种重金属浓度的背景值数据，对城区的生活区、工业区、山区、主干道路及公园绿地区的重金属综合污染程度做了评估，其污染浓度分别为：中度污染，严重污染，轻度污染，严重污染，中度污染，严重污染。

第二问运用 Spearman 秩相关系数模型对不同重金属元素浓度间的相关程度与功能区之间相关程度分别做了探究，并结合 8 种重金属污染物来源的相关资料，给出造成 8 种重金属污染的主要的可能原因，如：重金属 Ni 主要来自交通污染。

第三问，为定位造成重金属污染的源，本文先对不同重金属在该城区的扩散特征做了探讨，并详细探讨了如何定位污染源的方法：基于 K-means 聚类模型建立污染源位置定位模型。本文还就“是否存在不同重金属污染来自同一污染源”进行了讨论，运用模糊数学聚类模型给出了结果：存在 {Cd, Hg}，{Cr, Cu, Pb} 这两类集合内重金属元素来自同一污染源。最后求解出 8 种重金属各自的污染源坐标与同时造成 {Cd, Hg}，{Cr, Cu, Pb} 重金属污染的污染源坐标。通过与浓度分布图做对比，结果表明污染源坐标的确定是相当准确的。

第四问，通过研究如何量化城市土壤的自净化能力，引入与时间  $t$  相关的自净化数学模型，并对求解过程参数的转换做了详细分析，最后得出在问题所提供的数据基础上，还需要增加土壤的酸碱度测量值。

文末对文中解决各个问题所建的数学模型，从可操作性，精度，模型本身性质，原始数据带来的影响等角度进行模型优缺点的评价。最后做了该城区的环境分析报告，总结本文对该城区表层土壤重金属浓度问题的求解结果，并给出了有针对性的建议。

**关键词：**定性定量 综合评价 K-mean 聚类定位 土壤自净能力 Matlab （选取的方法我们也在写作课上介绍了，从研究问题、观察角度、解决方法方面选，耗时 5-10 分钟）

## 一 问题提出

### 1.1 题目背景与研究意义

此处最好花 30-40 分钟, 抄改 2-3 段约为 1/3 页的全文总起段。体现你对题目的初步理解。和“问题重述”共存 1 页。

### 1.2 问题重述

问题重述花 30-40 分钟，稍微改写下。如果不愿话费时间，则把原问题复制后，提炼成表格。即：原问题+自己提炼的要解决的问题点表格（表格内用简短的语句，切忌长句）

注意表格哦~

|       | ① | ② | ③ | ④ |
|-------|---|---|---|---|
| 提炼的条件 |   |   |   |   |
| 提炼的问题 |   |   |   |   |

## 二 问题分析

总起段。约 3-4 行，对拟问题做个分析方向的介绍。可讲你们想在哪些方面进行测试等。或 1-2 行，如“此部分，对本文对要解决的问题进行建模前的初步分析准备，以求为后文建模部分提供逻辑结构”。

### 2.1 基本思路

对“1.2 问题重述”中提炼的问题，逐条进行思路分析。对于复杂问题的思路约为 3-4 行，简单的问题不超过 3 行。

拟解决的问题①：特征，难点，可尝试的方向。

拟解决的问题②：

拟解决的问题③：

拟解决的问题④：

### 2.2 具体分析

#### 2.2.1 相关概念的定义

对你所不理解的某些特定词汇或概念给出准确的解释，不超过 5 个，每个约 1 行。如“高程数据”“污染源”。

(1)

(2)

#### 2.2.2 刻画不同重金属元素的空间分布（拟解决问题稍做修改，或不修改为标题）

2.2.2-2.2.6，按需删除和增加新的小节。此处可以介绍基本思路中你们欲详细开展的研究方法、步骤 以及可能存在的问题（如：模型阈值设置，概率模型模拟时的随机数据特征。）

#### 2.2.3 城区内不同区域重金属污染程度的确定

。

#### 2.2.4 导致重金属污染的原因

。

#### 2.2.5 重金属污染物扩散传播特征

。

#### 2.2.6 定位污染源位置

。

## 三 基本假设

1、假设：选取你们团队讨论中，投票后认为重要的假设。

理由： 这个必须有！有了你们论文总分+3 分。对你们的假设做出 1-2 行的解释，告诉大家，该假设对所解决问题理论模型的重要性。

2、假设暂不考虑影响土壤重金属形态分布的物理性质。

理由：影响城市重金属分布的因素有很多，主要有两大方面，一为土壤的理化性质

如 PH 值、土壤质地、胶体含量、离子含量、营养元素等；二为人类活动如添加到城市土壤中的重金属的数量、种类的影响这些相关性的因素都会影响土壤重金属形态分布，相同的土壤条件下，同种重金属添加数量不同重金属形态分布也不同。

#### 四 符号说明

| 符号            | 意义                       |
|---------------|--------------------------|
| $x,y$         | 该城区重金属采样点的坐标             |
| $z$           | 该城区重金属采样点的海拔高度           |
| $SSE$         | 误差平方和                    |
| $R^2$         | 判定系数                     |
| $DFE$         | 误差自由度                    |
| $RMSE$        | 均方根误差                    |
| $\bar{a}$     | 样本各重金属浓度的平均值             |
| $S$           | 样本各重金属浓度的标准差             |
| $P_i$         | 土壤中第 $i$ 种污染物的单因子污染程度指数  |
| $C_i$         | 第 $i$ 种污染物的实测值           |
| $S_i$         | 第 $i$ 种污染物的评价标准值         |
| $CV$          | 标准差系数                    |
| $(P_i)_{ave}$ | 土壤中第 $i$ 种污染物的单项污染指数的平均值 |
| $(P_i)_{max}$ | 土壤中第 $i$ 种污染物的单项污染指数的最大值 |
| $P_{综}$       | 内梅罗综合污染指数                |
| $\rho_s$      | Spearman 秩相关系数           |

## 五 模型的建立和求解

该章节大概页数为 12-17 页（总页数 20-30 页）

### 5.1 模型的准备

#### 5.1.1 研究重金属浓度空间分布时涉及到的概念说明

（此段概念说明，如在二问题中的笔墨较多，则可有可无。）

相关研究显示，城市土壤重金属污染分布受人类影响明显，城市土壤重金属污染有独特而复杂的空间分布特征，从人类活动密集的城市中心到市郊重金属含量有放射形递减的趋势。城市土壤重金属含量高于郊区及农田土壤的含量，重金属总量由城区到郊区再到农区逐渐减少，尤其是 Pb、Cu、Ni、Cr 主要污染离子变化明显。

Biharmonic 样条曲面插值：

可实现空间映射，采用后向映射和双线性插值进行图像重采样。在图像变形基础上结合图像融合技术实现了图像渐变。该方法使用灵活，操作者可以简单地改变特征点的数量和位置以控制变形细节而无需其他操作。实验结果表明，该方法产生的过渡图像平滑、自然，每幅变形图像边界稳定且有良好的可视性。

#### 5.1.2 数据的初步处理

（如有数据处理，必须有！）

数据处理，可以有筛选清洗、分类、简单的描述统计分析等。

为方便接下来各模型的选择、建立、计算与分析，故先针对题目所给出的数据做可预料的初步数据分类、统计描述、城区地形图层绘制等。

##### (1) 数据汇总分类：

- ① 将附件 1 与附件 2 参数汇总。
- ② 按相同功能区将总体数据分类。

##### (2) 数据的统计描述：

表 1 8 种元素的数据统计（上过写作课的同学别忘了，这个可是自动表注哦）

| 元素   | As                  | Cd                | Cr                  | Cu                  | Hg                | Ni                  | Pb                  | Zn                  |
|------|---------------------|-------------------|---------------------|---------------------|-------------------|---------------------|---------------------|---------------------|
| 统计量  | ( $\mu\text{g/g}$ ) | ( $\text{ng/g}$ ) | ( $\mu\text{g/g}$ ) | ( $\mu\text{g/g}$ ) | ( $\text{ng/g}$ ) | ( $\mu\text{g/g}$ ) | ( $\mu\text{g/g}$ ) | ( $\mu\text{g/g}$ ) |
| 平均值  | 5.676489            | 302.3962          | 53.50966            | 55.01674            | 299.7113          | 17.26185            | 61.74094            | 201.2026            |
| 标准差  | 3.024292            | 224.9876          | 70.00179            | 162.9151            | 1629.54           | 9.941423            | 50.05776            | 339.2325            |
| 标准误差 | 0.169328            | 12.59689          | 3.919348            | 9.121494            | 91.23671          | 0.556613            | 2.802696            | 18.99338            |

题目所给城区数据海拔高度的平均值为 42.46m，根据地理知识：海拔 0—200 米的广大地区为平原，不妨先猜想研究的城区处于我国某平原地区。

##### (3) 城区的地形图绘制

若所解决问题需要绘图，则最好做初步的数据成像处理。

为了之后更加深入的探究问题与方便直观上了解所研究地区的地形概貌，对题目所给采样点数据的坐标、海拔高度，在 Matlab 2011a 下使用其内置的 Biharmonic (v4) 样条差值算法进行三维坐标的插值处理，模拟该城区的三维立体图，如图 1 所示：

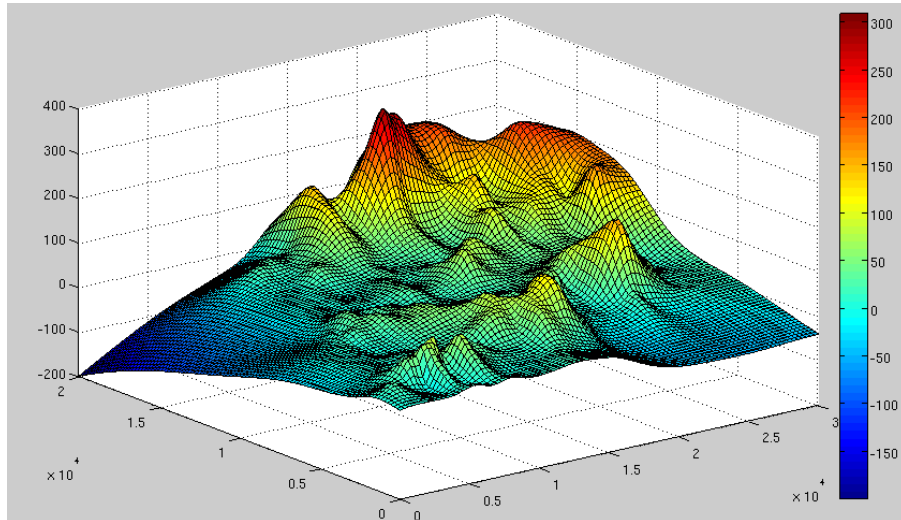


图 1 城区的三维地形图注意数字标签与图注、表注中文字要空一格，并使用自动表注功能  
同时，绘制了反应所研究城区整体地形的等高线图（图 2）：

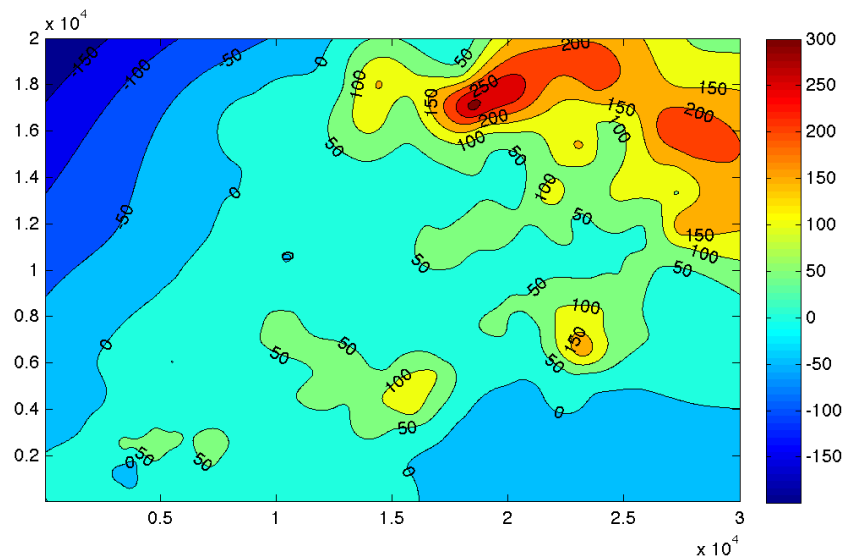


图 2 所研究城区的等高线地形图



## 5.2 定性刻画重金属浓度的空间分布

### 5.2.1 地形图上 8 种重金属元素浓度分布

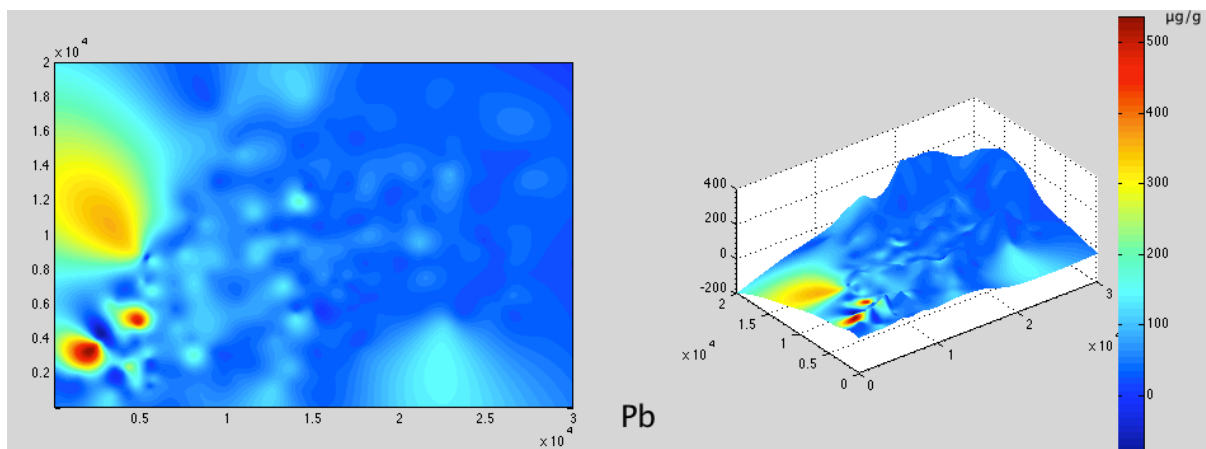


图 3 Pb 浓度的空间分布

图 3（如果你把图 2 删掉，选中“图 3 Pb 浓度的空间分布 图 3”后按 F9，这个可是会自动修改编号的哦！方法参加了我们课程的同学一定知道~）城区中 Pb 的浓度最大值达到  $472.48\mu\text{g/g}$ ，最小值为  $19.68\mu\text{g/g}$ ，平均值为  $61.74\mu\text{g/g}$ ，Pb 的背景值为  $31\mu\text{g/g}$ 。

### 5.2.2 重金属浓度空间分布的简要分析

## 5.3 定量刻画重金属浓度的空间分布

正文，总分形式写。这个是一个小的总起段。

考虑寻找另一种较为可行的且方便的量化方法，出于在不失准确度且简便操作的考虑，可以探讨：将三维转化成二维多项式拟合，将“坐标与浓度关系刻画”换成“海拔高度与浓度关系的刻画”。

### 5.3.1 重金属浓度数据与海拔高度关系的探讨

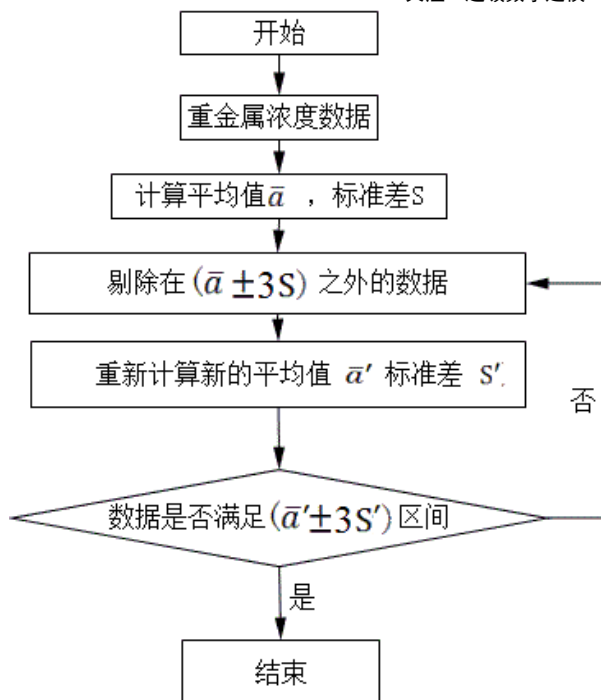
为能实现在二维上研究量化重金属浓度的空间分布，故需探讨重金属浓度数据与海拔高度间的关系。

（1）剔除城区的重金属浓度数据中的“异常值”

同时，由于在该城区重金属浓度数据中存在个别浓度值极大的数据（重金属浓度极其严重的部分区域），其对后续的插值会造成较大的影响，它的存在会造成连续表面的中断，直接影响变量的分布特征，使半方差函数失去结构性，所以需要对相关重金属浓度数据进行预处理。暂且不考虑城区中重金属浓度极其严重的部分区域，剔除这些对样本数据影响非常大的区域的取样数据，从而刻画该城区排除个别污染非常严重区域外，城区的土壤重金属浓度的空间分布。不妨将剔除的少量数据点称其为“异常点”。

本文采用域法识别特异值，即样本平均值  $\bar{a}$  加减三倍标准差  $S$ ，在此区间  $(\bar{a} \pm 3S)$  以外的数据均定为特异值。

剔除异常点的流程图为：（考考你们，写作课上，我们说这个图用什么软件画？）



## (2) 探讨金属浓度与海拔高度的联系

在现实生活中随着海拔高度的增加，人类生活足迹对自然的影响大体呈递减趋势。如：我国城市与人口数量分布在地形三大阶梯上的变化。

为了探讨重金属浓度与海拔间的联系，结合该城区的重金属浓度数据的散点分布图（使用 Excel2010 制图），发现的确有较为明显聚合的趋势。

理由是：样本取样时是以所考察的城区划分间距 1 公里左右的网格子区域后，在每平方公里内取 1 个采样点的采样方法采样，由于该地区的平均海拔为 42.46m，故不难预料取值点多落在低海拔地段。在此以 As 浓度与海拔关系为例，进行说明。“As 浓度与海拔”的散点分布图见附录。

分析可得该城区由于地形缘故，整体上 As 污染浓度与海拔的高度呈现较为明显的聚集关系。即：As 污染点聚集在海拔较低的地区，而海拔较高的地区重金属污染点较少。为了量化分析重金属浓度空间分布关系，以求能方便反应该城市整体 As 污染浓度与海拔高度间关系的角度出发，可采取分区间探讨法。

通过分析统计 8 种重金属数据在不同海拔区间上的频数，最终确定了频数分布较均匀的 13 个海拔区间（单位：米）的分类：

## 5.4 不同功能区重金属污染程度的综合评定

所有的正文部分，按照写作培训课说的“总起”方式来写。段落间某段总。一段中，首句或末句总。

由于污染程度是一个较为抽象的概念，虽然已经有了各个区域取样的数据，但区域的污染程度的界定涉及到对抽象概念的定量评价。就好像已经拥有一群人的身高、体重数据，但对于单个人的“高矮胖瘦”则不好直接判定。故对抽象概念进行定量评价时，需要寻找相对稳定的对比标准。

### 5.4.1 与土壤重金属污染评价相关的研究

目前土壤重金属评价方法包括生态危害指数法、地积累指数法、单因子指数法和内梅罗综合污染指数法、模糊贴近度方法、潜在生态危害指数法等等。这些评价方法在评价土壤重金属污染程度上都有所运用，正如其评价方法一样，其评价标准也多种多样。

但总体而一言还是缺乏一个比较系统而又能客观评价土壤重金属污染程度的标准体系，这也主要是因为土壤，特别是城市土壤重金属污染的研究刚刚起步，还有待更深入的研究和发展。

针对所要解决的问题本身来说，对某个“区域的污染程度”的综合评价，基础数据涉及多种主要影响因素（8 种重金属），且题目本身提供了该城区的土壤中 8 种重金属元素的背景值。考虑到应用上的方便，故本文采用目前最普遍的内梅罗综合污染指数法来考究土壤重金属污染的程度。

#### 5.4.2 对 N. L. Nemerow 指数的说明

N. L. Nemerow（内梅罗）指数 内梅罗指数是一种兼顾极值或称突出最大值的计权型多因子环境质量指数内梅罗指数的基本计算式为：

$$I = \{[(\max I_i)^2 + (\text{ave} I_i)^2] / 2\}^{1/2} \quad (3-7)$$

式中  $\max I_i$  为各单因子环境质量指数中最大者， $\text{ave} I_i$  为各单因子环境质量指数的平均值。内梅罗指数特别考虑了污染最严重的因子，内梅罗环境质量指数在加权过程中避免了权系数中主观因素的影响，是目前仍然应用较多的一种环境质量指数。

#### 5.4.3 对所给背景值数据的说明

题目所给出的该城区的土壤中 8 种重金属元素的背景值，其实是为解决综合评价污染程度所提供的，在土壤重金属研究中最常用到的参考标准。

背景值是指在不受污染的情况下，各化学元素在土壤中的正常含量。从本质上讲，“不受污染源明显影响”只是一个相对概念。土壤成因和类型不同以及人类活动影响必然导致土壤化合物含量的空间差异，因此土壤背景值是个范围值，不是个确定值。

#### 5.4.4 基于 N. L. Nemerow 指数法重金属污染程度评价模型的建立

对抽象概念进行定量评价时，需要寻找一个对比标准。则可采用该城区元素背景值为基准，先量化单个因子（单种重金属元素浓度）对污染程度指标，而后再结合 8 种重金属元素分别的污染程度指标，划分综合污染等级，进行定量的综合评价。

本文使用单重金属元素的浓度为考核污染程度的指标：单因子污染指数；用此来判断所研究区域表层土壤中该种重金属的人为污染情况。同时通过分析各元素的变异系数来衡量各单因子观测值的变异程度，即单因子污染程度数据间的差距大小。结合单因子污染指数，计算各个区域的 N. L. Nemerow 综合评价指数，综合表征每个区域的土壤污染情况。

##### 5.4.4.1 单因子污染指标、重金属元素污染情况分级标准及变异系数的确定

#### 5.4.5 不同区域重金属污染程度评价模型的求解

由表格 7 所得出的不同功能区的重金属污染等级评估中显示：生活区中 Cu、Zn 污染严重化；工业区除 As、Cr、Ni 外其余重金属浓度均达到严重等级；山区的重金属污染水平都属于轻度；城市主干道受 Cu、Hg、Zn 重金属污染严重化；公园绿化地区只有 Hg 的污染达到了严重程度。总的来说该城区整体上 Cu、Hg 污染严重化，整体城市重金属综合污染等级达到严重等级。

## 5.5 探究该城区重金属污染的主要原因

### 5.5.1 研究城市重金属主要原因的方法论

该问由重金属污染浓度的结果去考究导致的原因，即“溯源”，这样对于“源”的可能性相当多。一一研究显然不现实，故考虑从我国导致重金属污染的普遍原因（即主要的人类活动）下手，借此为研究该城市重金属污染的主要原因提供依据。这类似于现实生活中，我们看到一个肥胖的人时，脑中浮现的是大部分人由于饮食偏好导致肥胖，故饮食偏好可作为评判此人肥胖原因的主要参考。

作为该问题本身提供的数据量来说，对主要原因的探讨结果，并不能作为十分肯定的依据只能作为参考。具体的原因应该获取到与此地相关的更为丰富的信息，如：工业区类型，工业区数量或工厂的类型等，在此基础上才能做到更为精准的原因定位。

### 5.5.2 探讨重金属元素浓度间的相关程度

探讨理由：城市土壤重金属的来源存在两种途径：一种是来自于成土母质；另一种是外源输入，来自于人类各种生产活动。在现实中人类的生产活动，给自然土壤带来各种各样的污染物质，使其中所含的重金属元素随着污染物质的排放而在土壤中富集，而且不同重金属元素的富集可能来自同一污染源。N 类重金属元素浓度之间的相关程度高，则表明有大的可能性这 N 种重金属的来源是相同。

因为重金属污染物的来源多种多样，所以各种重金属元素之间的影响普遍情况下并非呈线性关系，或线性关系不明显。在此，可只针对重金属元素之间的趋势（伴随增加或减少的性质）做分析。

### 5.5.3 建立 Spearman 秩相关系数模型分析各重金属浓度数据间的相关程度

#### 5.5.3.1 Spearman 秩相关系数介绍

Spearman 秩相关系数是一个非参数性质（与分布无关）的秩统计参数，由 Spearman 在 1904 年提出，用来度量两个变量之间联系的强弱 (Lehmann and D' Abrera 1998)。Spearman 秩相关系数可以用于 R 检验，同样可以在数据的分布使得 Pearson 线性相关系数不能用来描述或是用来描述或导致错误的结论时，作为变量之间单调联系强弱的度量。只要在 X 和 Y 具有单调的函数关系的关系，那么 X 和 Y 就是完全 Spearman 相关的，这与 Pearson 相关性不同，后者只有在变量之间具有线性关系时才是完全相关的。

我国的环保总局也在《环境质量报告书编写技术规定》中推荐使用 Spearman 秩相关系数法对污染变化趋势进行定量分析。

#### 5.5.3.2 Spearman 秩相关系数计算原理

假设原始的数据  $x_i, y_i$  已经按从大到小的顺序排列，记  $x'_i, y'_i$  为  $x_i, y_i$  在排列后数据所在的位置，则  $x'_i, y'_i$  称为变量  $x_i, y_i$  的秩次，则  $d_i = x'_i - y'_i$  为  $x_i, y_i$  的秩次之差。

如果没有相同的秩次，则  $\rho_s$  可由下式计算

$$\rho_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$$

如果有相同的秩次存在，那么就需要计算秩次之间的 Pearson 的线性相关系数

$$\rho_s = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2 \sum_i (y_i - \bar{y})^2}}$$

一个相同的值在一列数据中必须有相同的秩次，那么在计算中采用的秩次就是数值在按从大到小排列时所在位置的平均值。

Spearman 秩相关系数的符号表示  $X$  和  $Y$  之间联系的方向。Spearman 秩相关系数是正，则  $Y$  随着  $X$  的增加而增加；若 Spearman 秩相关系数就是负的，则  $Y$  随着  $X$  的增加而减小。若 Spearman 秩相关系数为 0，表示随着  $X$  的增加， $Y$  没有增大或减小的趋势。

随着  $X$  和  $Y$  越来越接近严格单调的函数关系，Spearman 秩相关系数在数值上越来越大。 $\rho_s \in [-1,1]$ ，当  $X$ 、 $Y$  有严格单增的关系时， $\rho_s = 1$ ；当  $X$ 、 $Y$  有严格单减的关系时， $\rho_s = -1$ 。

## 5.5.4 重金属间相关程度的求解结果

结果分析要“长”!!!! 尽可能对结果分析要详细。结果说明了什么，特征描述，让评卷人关注什么值，为什么等。1 页结果分析都不为过！

由于题目样本数目多于 30 个，且采样点的选取，按照间隔 1 公里的原则进行采样，因此结合采样点所在功能区不同，随机选取 5 类功能区中 6 个采样点，作为重金属之间 Spearman 秩相关系数的计算样本点。

随机选取出两组样本数目为 30 的采样点集，用 MATLAB 计算两个采样点集 8 种重金属元素之间的 Spearman 秩相关系数，如附录表 9 所示。从表 9 可以看出，采样点集 1 中，Cu-Hg 的 Spearman 秩相关系数为 0.345 ( $>0.306$ )，Cd-Zn 的 Spearman 秩相关系数为 0.326 ( $>0.306$ )。由此说明，在显著水平  $\alpha = 0.05$  的条件下，Cu 的含量与 Hg 的含量相关，随着彼此的增加而增加；Cd 的含量与 Zn 的含量相关，随着彼此的增加而增加。

从表 10 可以看出，采样点集 1 中，As-Cd 的 Spearman 秩相关系数为 0.328 ( $>0.306$ )，Cr-Hg 的 Spearman 秩相关系数为 0.444 ( $>0.432$ )。由此说明，在显著水平  $\alpha = 0.05$  的条件下，As 的含量与 Cd 的含量相关，随着彼此的增加而增加；在显著水平  $\alpha = 0.01$  的条件下 Cr 的含量与 Hg 的含量相关，随着彼此的增加而增加。

按城区功能区域划分的不同，用 MATLAB 计算各个功能区重金属含量之间的 Spearman 秩相关系数，由表 11 和表 12 可以看出，两组随机选取的采样点集，计算得出的各个功能区重金属含量之间的 Spearman 秩相关系数完全一样，由此说明本题的 319 个采样点的选取具有随机性，从全部采样点中随机选取 30 个不同的样本构成 Spearman 秩相关系数的计算样本，计算得出的 8 种重金属之间 Spearman 秩相关系数虽有不同情况存在，但从性质上而言，其相关性仍存在，即认为上述两组随机采样点集计算得出的 Cu-Hg、Cd-Zn、As-Cd、Cr-Hg 的相关性可信。

生活区与山区、交通区、公园绿地区之间的 Spearman 秩相关系数分别为 0.976、0.333、1，山区与交通区、绿地工业区之间的 Spearman 秩相关系数分别为 0.357、0.97，交通区与公园绿地区之间的 Spearman 秩相关系数为 0.333。说明，在显著水平  $\alpha = 0.05$  的条件下，三个区域两两之间存在相关性。

工业区与交通区的 Spearman 秩相关系数为 0.786，在显著水平  $\alpha = 0.05$  和  $\alpha = 0.01$  的条件下，均可信，即两个区域之间存在相关性。

上述数据说明，各个功能区之间的重金属污染程度随人类活动的关联性而息息相关。除工业区外，其他四个功能区间两两相关，而工业区主要与交通区相关，这是由其工业区经常与交通区毗邻所造成的。

## 5.5.5 可能造成重金属污染原因讨论

查阅造成重金属污染的原因的有关资料，可以了解到：

目前，工业能源大多以石油、煤、天然气等化石燃料为主，这些化石燃料中含有大量重金属和类金属，如 Hg、Cr、Cd、As、Pb 等元素；

90 年代至今，对交通重金属复合污染研究成为了学界研究的重点. 现在的观点普遍认为 Cd、Ni、Pb、Zn 随距离公路的距离增加呈下降趋势；

Zn、Pb、Cu 这三种重金属污染与机动车尾气排放有关；

污水灌溉已经成为城市农田重金属的最主要来源，这导致土壤重金属 Hg、Cd、Cr、AS、Cu、Zn、Pb 等元素含量的增加。

结合上述文献资料，分析本题该城区重金属污染的主要原因，可以看出重金属元素之间的相关性与各自产生的人类活动来源之间的对应关系，计算得到的 Spearman 秩相关系数反映出相关性质与文献资料介绍情况相符。

综上所述，重金属污染的主要原因应为：

## 5.6 寻找重金属污染源的地理位置

### 5.6.1 8 种重金属的扩散特征讨论

借助“5.2 定性刻画金属浓度的空间分布”已做的 8 种重金属元素浓度的空间分布图（见附录三），初步探讨扩散规律：

将 8 种重金属元素浓度平面模拟图汇总，与发现 8 种重金属元素的扩散特征都可描述为：以某点为中心向四周辐射，随范围的增大浓度降低。

考虑到八种重金属元素本身物理、化学特性都存在较大差异。逐个探讨重金属元素的具体扩散特征，不同污染源相同重金属元素扩散复合特征显得繁琐而低效。并且即使讨论出具体的扩散特征，由于重金属的扩散与海拔高度的变化间又存在一定的联系，扩散特征又得引入海拔高度重新讨论，更显得复杂与不可操作性。

在这种情况下，鉴于接下来所建的污染源定位模型需要在保证精度的情况下又具简便的可操作性，故对重金属的扩散形式与具体特征不做深入的探究，只需为接下来污染源的定位提供参考。

在探讨重金属扩散过程与寻找污染源的过程中，可采用聚类分析的方法，于是引入此种方法来定位重金属污染源的地理位置。思路如下：

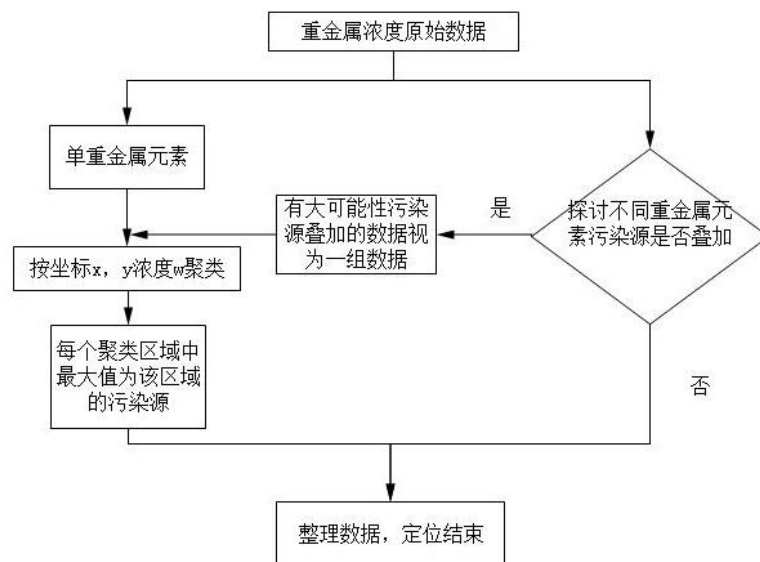


图 1 由讨论扩散特征得出定位污染源的思路

### 5.6.2 基于 K-means 聚类模型建立污染源位置定位模型

模型部分很多是既定的。可以借鉴你们提前收集的算法资料。

在聚类模型的选择上，选择了经典的 K 均值聚类模型。



K-means 算法是典型的局域原型的目标函数聚类方法的代表，它是数据点到原型的某种距离作为优化的目标函数，利用函数求极值的方法得到迭代运算的调整。算法采用误差平方和准则函数作为聚类准则函数。

K-means 算法的主要思想是：通过迭代过程把数据集划分为不同的类别，使得评价聚类性能的准则函数达到最优，从而使生成的每个聚类区类内紧凑，类间独立，即，同一类中的个体具有高度的同质性，不同类之间的个体具有高度的异质性。

考虑采样点的在 0-xy 平面坐标系中的坐标点(x, y)和重金属元素浓度这三个指标，从三维空间进行聚类。由分类情况得出每一类中浓度最高点，即得出污染源的位置。

K-means 模型计算流程图，如下：

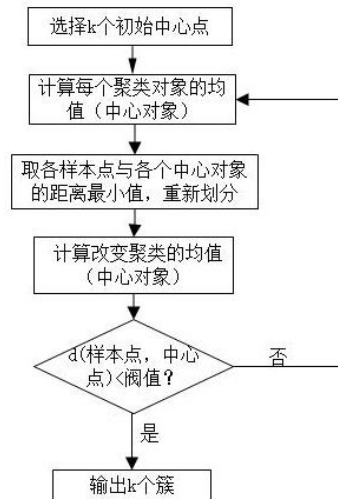


图 2 K 均值聚类的算法流程

实现 K-means 算法具体步骤如下：

- STEP 1: 为每个聚类确定一个初始聚类中心，有 k 个初始聚类中心；
- STEP 2: 将样本集中的样本按照各项指标的最小欧式距离原则分配到最邻近聚类；
- STEP 3: 使用每个聚类中的样本均值作为新的聚类中心；
- STEP 4: 重复步骤 2、3 直到聚类中心不再变化；
- STEP 5: 结束，得到 K 个聚类
- STEP 6: 输出：k 个簇，使平方误差准则最小。

其中，不断将样本分配给距离它们最近的中心向量，并使目标函数

$$\sum_{i=1}^n \min_{j \in \{1, 2, \dots, k\}} \|x_i - p_j\|^2$$

的值减小。

更新簇平均值  $\bar{x}_i$ ， $c[i]=\{\text{所有标记为 } i \text{ 的 data}[j]\text{-之和}\}/\text{标记为 } i \text{ 的个数}$ ：

$$\bar{x}_i = \frac{1}{|C_i|} \sum_{x \in C_i} x$$

计算准则函数 E:

$$E = \sum_{i=1}^k \sum_{x \in C_i} |x - \bar{x}_i|^2$$

### 5.6.3 对是否存在不同重金属污染来自同一污染源的探究

#### 5.6.3.1 运用模糊聚类分析法对八种重金属进行相关性归类

对所使用的方法要做下简单介绍

为了确定是否存在不同重金属污染来自同一污染源的猜想。如成立，需将元素复合分析。

考虑采用模糊聚类分析对 8 种重金属元素进行分类，把相关性强的几种重金属归为一类，再根据模糊聚类情况，以每一个重金属分类为标准进行 K-means 聚类分析，确定出此类中所含的这些重金属复合污染物的污染源位置。

模糊聚类分析法模型的具体步骤：

##### (1) 重金属元素的模糊聚类分析

设  $A_j$  ( $j=1,2, \dots, 8$ ) 表示第  $j$  个重金属元素在各个样本点的浓度信息集。

利用模糊数学建立隶属函数，在此选取各个重金属元素在各个样本点的测量浓度作为隶属度函数：

$$\mu_{A_j}(x) = e^{-\left(\frac{x-a_j}{b_j}\right)^2}$$

$$\text{式中, } a_j = \frac{\sum_{i=1}^{319} a_{ij}}{319}, \quad b_j = \sqrt{\frac{1}{318} \sum_{i=1}^{319} (a_{ij} - a_j)^2}, \quad (j=1,2, \dots, 8)$$

利用格贴近度建立模糊相似矩阵：

令模糊相似系数

$$r_{ij} = e^{-\left(\frac{a_j - a_i}{b_i + b_j}\right)^2} \quad (i,j=1, 2, \dots, 8)$$

则模糊相似矩阵

$$R = (r_{ij})_{8 \times 8} \quad (i,j=1, 2, \dots, 8)$$

根据模糊数学知识,  $R$  为模糊相似矩阵, 阶数为  $n$ , 则存在一个最小的自然数  $k$  ( $k \leq n$ ), 使得  $R^k$  为模糊等价矩阵 (若  $R$  为模糊等价矩阵, 则  $r_{ii} = 1, r_{ij} = r_{ji}$ , 且  $R^2 \leq R$ ), 且对一切大于  $k$  的自然数  $l$ , 恒有  $R^l = R^k$ 。  $R^k$  称为  $R$  的传递闭包矩阵, 记为  $t(R)$ 。

从  $n$  阶模糊相似矩阵  $R$  出发, 依次求平方:

$$R \rightarrow R^2 \rightarrow R^4 \rightarrow \dots$$

直到

$$R^{2^i} \circ R^{2^i} = R^{2^i} \quad (2^i \geq n, i \geq \log_2 n)$$

为止, 则

$$t(R) = R^{2^i}$$

其中, “ $\circ$ ” 为模糊运算符。

设  $A = (a_{ik})_{m \times s}$ ,  $B = (b_{kj})_{s \times n}$ , 定义



$$A \circ B = (c_{ij})_{m \times n}$$

$$c_{ij} = \max\{(a_{ik} \wedge b_{kj}) | 1 \leq k \leq s\}$$

其中，“ $\wedge$ ”为取大值运算。

求得  $R$  的传递闭包  $t(R) = R^k$ ，即得模糊等价矩阵  $R^k$ 。然后，由大到小取一组  $\lambda \in [0, 1]$ ，确定相应的  $\lambda$  截矩阵，则可以将  $R^k$  分类。 $R_\lambda^k$  指  $R^k$  中各元素中大于  $\lambda$  的数用 1 代替，小于  $\lambda$  的数用 0 代替，重新得到的  $n$  阶方阵。

分类方法为，确定初始  $\lambda$  值，这里选取初值  $\lambda = 0.998$ ， $R_\lambda^k$  中取出相异的列向量，如果  $R_\lambda^k$  中的元素  $V_{ij} = 1$ ，则将集合  $x$  中的元素聚为一类。各个列向量中 1 所对应的因子合成一类，有几个相异的列向量就可分成几类。

### 5.6.3.2 “不同重金属元素同一污染源猜想”的探究结果

结果的呈现，在正文的多个地方要出现。即你要针对每个问题做详细的结果分析。运用 Matlab2011a 进行 8 类重金属元素的模糊聚类分析编写程序求解，代码见附录，当  $\lambda=0.998$  时，

$$R_\lambda = \begin{matrix} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \begin{matrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{matrix} \end{matrix}$$

此时，得到的分类为： $\{As\}$ ， $\{Cd, Hg\}$ ， $\{Cr, Cu, Pb\}$ ， $\{Ni\}$ ， $\{Zn\}$ 。故证明了的确实是存在不同重金属来自同一污染源的猜想。在使用 K-means 进行污染源分类时，还需结合考虑  $\{Cd, Hg\}$ ， $\{Cr, Cu, Pb\}$  这两类重金属元素来自同一污染源这样的情况。

### 5.6.4 污染源位置定位结果

使用 Matlab2011a 对基于 K-means 聚类模型定位污染源进行求解，其中 K 值的确定为结合重金属本身的浓度扩散图示（见附录五：K-mean），多次选择比较后确定：

具体污染源位置坐标求解结果见附录，此处对同一重金属元素不同污染程度下污染源个数的统计：

## 六 城市的演变模式研究

如有时间，自己想想此次选中的问题的研究背景，并对此做下分析。（此处需要整理文献比较多）为什么？因为你要表达：你们写的这个模型其实是短时间内看了超多文献，并在前人研究的基础上提出了自己新的改进或解决方案。

为试图为研究城市化演变过程模式，提出有建设性、合理性、实用性的研究城市地质环境模型，我们考虑先对该门学科的发展程度做了解，深化认识后，再开始尝试建立新的对城市化演变模式下地质的变化模型。

方法论：

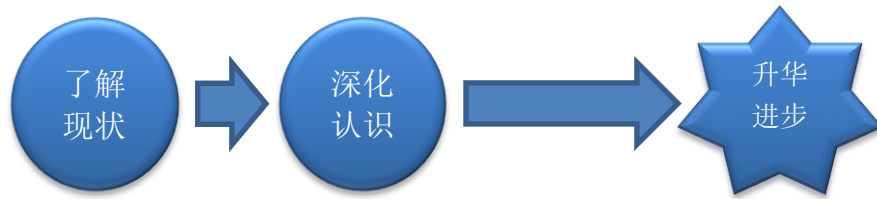


图 3 升华对城市演变模式的研究的方法论

## 6.1 城市地质环境的相关研究

“城市地质学”这一概念最早是由 R. F. Iegget 提出的，是一门关于在城市及周边地区或潜在城市化地区的特定空间范围内，综合考虑各种地质要素，研究这些要素为城市发展所提供的资源、施加的约束调价和城市对各种地质要素所产生的影响，为城市规划、建设和管理服务的学科。

### 1. 国外近 10 年来城市地质学研究进展

纵观近 10 年来国外城市地质学的研究进展，边坡失稳、地下水污染、土壤污染等城市地质环境问题分析与评价仍是城市地质学的焦点，城市地质如何与城市规划结合起来是城市地质学的热点问题。城市三维地质建模、城市地质灾害风险管理等则是城市地质学的前沿问题。利用现代信息技术和手段、以 GIS 为平台的数字化空间数据库和计算机成图，已成为现代城市地质工作行测主流工作模式。

### 2. 我国城市地质工作研究进展

近十年来，国内城市地质学研究主要集中在城市地质环境方面。例如：加强地质环境评价方法的研究、城市环境地质研究、地球化学、地球物理相结合，把研究领域拓展到城市群。

在城市地质环境评价方面，城市工程地质环境评价、地下水资源脆弱性评价等单项地质环境评价比较多，综合性的评价研究少。关于人类活动对地质环境的影响方面研究不多，有的城市地质环境评价没有很好地考虑人类活动因素，仅仅使用“人口密度”“单位面积 GDP”“工程活动强度”这样很粗略的指标。虽然 GIS 在城市地质环境评价领域已经得到了应用，但缺乏专门的城市地质环境评价系统。

在中国地质调查局的推动下，城市地质调查方法、城市地质立体填图技术等方面的研究逐渐兴起。计算机技术在地质调查、地质填图方的应用越来越广泛、越来越深入，三维地质建模技术成为研究的热门领域。

## 6.2 城市环境与城市演变关系过程中对城市土壤的自净能力的探讨

方程的图形是一条有单峰的曲线，随着时间  $t$  的增大，有效污染物质的含量逐渐趋于平衡位置  $Y_0$ ，说明自净能力强的情况下土壤环境的恢复快。

(1) 自净作用等于临界值情形（临街阻尼）： $n=k$

$$y = e^{-\pi t}(c_1 + c_2 t)$$

则：

$$Y = Y_0 + e^{-\pi t}(c_1 + c_2 t)$$

式中  $c_1, c_2$  为待定参数。

由于  $\lim_{t \rightarrow \infty} t e^{-t}$ ，从而可以看出，随着时间  $t$  的增大，污染物质的含量将会趋向于平衡位置。

有上述分析可知，当土壤自净作用弱时，土壤有效污染物的量的函数曲线类似于阻尼振荡；当土壤自净作用强时，土壤有效污染物的函数曲线只有一个峰值，无振荡运动，说明土壤恢复快。

其中的  $Y_0$  和  $Z_0$  是需要我们测出的，而待定系数  $c_1, c_3$  可利用我们的在上述模型中得到的污染量算得。

定义有效污染量

$$C_{pH} = \frac{\text{总量}}{(pH/7)^2}$$

由上分析，为研究城市土壤演变模式的土壤自净能力时，相对题目所给数据来说，还需要增加的测量信息为：土壤的 PH 值。

## 七 本文所建模型优缺点分析

此处优缺点分析，按自己先前第“一”章中，提炼的问题来做小节。然后优缺点的分析长度如下所示。最后你们的文字分布要看起来就像真正的红烧肉，肥而不腻。

### 7.1 重金属浓度的空间分布

定性分析较为直观的刻画了重金属浓度的空间分布，且现实意义；而定量分析首先尝试了对分布在二维空间的离散点进行三次样条差值后函数拟合，但拟合效果不佳，故排除了求其浓度的分布函数的方法的可行性。然后，采用域法识别特异值剔除浓度异常点，再分析重金属浓度与海拔间的联系，进行二维曲线拟合，拟合结果具有较高精度。

为了避免各种不知名的误差影响结果，我们剔除了部分对整体情况的分析有较大的影响的异常点，再进行二维曲线拟合量化，排除个别异常点对整体情况的影响使整体情况分析方便更具合理性。不足之处是无法完整衡量该城区的浓度分布。

### 7.2 不同功能区重金属污染程度的综合评定

采用单因子指数法，避免了污染物的自然来源的影响，能够更准确地判断出人类活动带来的污染影响。另外，采用内梅罗综合污染指数，是目前研究污染程度领域运用最为普遍的一项指标，具有计算简单和兼顾极值或称突出最大值的特点，在加权过程中避免了权系数中主观因素的影响。内梅罗综合污染指数整合各个单因子指数，即各种重金属的污染情况，综合评价不同功能区重金属污染程度，考虑全面，总体兼顾个体；它简单实现，避免主观因素的影响，多方面研究实例为其提供理论可信度，不足之处就是无法反映污染程度来源

### 7.3 探究该城区重金属污染的主要原因

利用重金属之间的相关性来分析各重金属污染来源的相同的情况，采用 Spearman 秩相关系数，它是一个非参数性质（与分布无关）的秩统计参数，只要变量之间存在单调函数关系，不管是否是线性关系。其优点有两点：一充分体现出变量之间的相关性质；二充分体现出变量之间的相关性质。不足之处为检验的临界值只有到样本数小于等于 30 的情况，对于样本数大于 30 的情况，没办法直接对全部样本进行计算检验。

### 7.4 寻找重金属污染源的地理位置

考虑污染源传播的辐射影响区域内，采样点的在 0-xy 平面坐标系中的坐标点  $(x, y)$  和重金属元素浓度这三个指标存在相似性，故从同一污染源辐射区内的相似性和不同污染源辐射区间的相异性，作为污染源位置确定的考察标准，建立 K-means 聚类模型。

此模型是典型的局域原型的目标函数聚类方法的代表通过迭代过程把数据集划分为不同的类别，使得评价聚类性能的准则函数达到最优，从而使生成的每个聚类区类内紧凑，类间独立，即，同一类中的个体具有高度的同质性，不同类之间的个体具有高度的异质性。其优点有三点：一为算法思想简单，容易理解，实现的难度也较小；二为在簇与簇之间区别明显时，分类效果较好；三点初始条件要求简单。不足之处：分类的结果会受聚类个数不同的影响，在无法确定类的数目情况下，需进行改进。

## 八 对该城市的环境报告

这个章节，可有可无。基本是用来表达你们有大把时间用来再对所解决的问题做个回顾。结合自己的模型结果对所研究问题本身做个评价。属于威慑评卷老师的部分，告诉他，你们不慌不忙的结束全文~

与自然生态系统相比，城市生态系统是社会—经济—自然相结合的复杂生态系统，也是一个极不稳定的生态系统。在这个生态系统中，城市土壤是其重要的组成部分，也是其食物链流动的物质基础。所以，土壤质量的优劣直接影响城市的生产、生活和发展，对城市的可持续发展有着重要的意义。

本文通过对已知数据资料进行数学模型的建立，研究重金属在该城区的空间分布，完成该城区的污染程度，分析重金属污染来源的可能性，并完成了对该城区污染源位置的定位。通过本文可以了解到该城区与表层土壤的重金属污染的相关情况：

(1) 该城市各个功能区的综合污染程度为：

| 功能区 | 污染综合等级 | 功能区    | 污染综合等级 |
|-----|--------|--------|--------|
| 生活区 | 中度     | 主干道路区  | 严重     |
| 工业区 | 严重     | 公园绿地区  | 中度     |
| 山区  | 轻度     | 城区总体情况 | 严重     |

从研究功能区的情况看：生活区中 Cu、Zn 污染严重化；工业区除 As、Cr、Ni 外其余重金属浓度均达到严重等级；山区的重金属污染水平都属于轻度；城市主干道受 Cu，Hg，Zn 重金属污染严重化；公园绿化地区只有 Hg 的污染达到了严重程度。总的来说该城区整体上 Cu，Hg 污染严重化，整体城市重金属综合污染等级达到严重等级

(2) 8 种重金属的含量相对背景值的增长速率：Hg>Cu>Zn>Cd>Pb>Cr>As>Ni，8 种重金属元素中相对该地背景值增长速率最大的为 Hg，其次为 Cu，分别为 756.31% 和 316.79%。

造成 8 种重金属污染的原因有：

| 元素种类 | 污染原因           | 元素种类 | 污染原因      |
|------|----------------|------|-----------|
| As   | 工业污染，农业污染      | Hg   | 工业污染，农业污染 |
| Cd   | 工业污染，交通污染，农业污染 | Ni   | 交通污染      |
| Cr   | 工业污染，农业污染      | Pb   | 工业污染，农业污染 |
| Cu   | 交通污染，农业污染      | Zn   | 交通污染，农业污染 |

(3) 造成该城市不同程度重金属污染的污染源数量：

| 程度\元素 | As | Cd | Cr | Cu | Hg | Ni | Pb | Zn |
|-------|----|----|----|----|----|----|----|----|
| 严重污染  | 7  | 27 | 11 | 35 | 27 | 1  | 22 | 26 |
| 中度污染  | 23 | 10 | 11 | 7  | 5  | 9  | 11 | 9  |
| 轻度污染  | 21 | 8  | 22 | 5  | 14 | 28 | 21 | 15 |
| 总计    | 51 | 45 | 44 | 47 | 46 | 38 | 54 | 50 |

其中{Cd, Hg}, {Cr, Cu, Pb}这两大类重金属元素是来自同一污染源的

20 世纪中、后期以来,世界的城市扩张急剧加速,城市人口快速增加,人类的生产 and 生活方式随之发生了重大变化,自然资源的开发和利用强度加大,并且由于诸多不科学的利用,资源 and 环境遭受着各种人为破坏。作为世界上人口最多、耕地后备资源匮乏的国家,我国的粮食安全,甚至政治稳定都依赖于土壤资源的合理利用和保护。但在我国社会经济发展都还处在初级阶段这样的客观背景下,生产发展具有迫切性,土地管理和保护工作相对滞后,土壤管理和利用不当的事件时有发生,不可避免地带来不同程度的土壤资源浪费和土壤质量退化。

“往者不可谏,来者犹可追”,认清现阶段城市化背景下人类活动强度和土壤演变状况,唤醒人类的环保意识,对于今后的土壤管理和保护具有重要作用。

综合该城区的污染情况,我们建议:

(1)、政府部门要加强对环境相关的法律法规的宣传,确保宣传到位,并普及公民的教育;

(2)、公民要认真学习科普知识,自觉接受教育,积极参与公益活动,努力提高自己的修养,并积极向周围的人传播相关的环境科学知识,尽自己最大努力置身于环境的改善;坚持从自我做起,从现在做起,努力提高自身的综合素质;

(3)、企业部门要自觉遵守我国环境保护法等,尽可能减少人为的污染;在追求利益的同时,不忘记保护环境,并且要努力把保护环境放在首要位置。

这个原文其实完稿时候 65 页~~然后最后花 3 小时删剩 29 页。

这个过程告诉我们,有得删的论文情况比没得删的要好得多得多。。。

## 参考文献

参考文献，回忆下写作课上，如何快速利用 google 学术获取标准的格式。

参考文献部分，我故意留下没有自动编号的形式，更深层的意思是。。告诉写作的那位，像我们写作课介绍的所有细节特征操作，其实都是要实现自动化!!! 编号自动化！算法流程自动化!!

- [1] 司守奎，数学建模算法与程序，海军航空工程学院，2011；
- [2] 居玲，研究生教育质量评价主体的问题与重建研究，西南大学学位论文，2008；
- [3] 李会军等，浅谈研究生学位论文开题质量监控评价体系，管理观察，397；25，2010；
- [4] 高嵩，基于证据理论的综合评比方法，华中科技大学硕士学位论文，2004；
- [5] 李鑫，济南市主要土壤类型在不同功能区的重金属形态分析，2008.5.17；12；
- [6] 汪权方、陈百明、李家永等.城市土壤研究进展与中国城市土壤生态保护研究[J].水土保持学报，2003，17(4):142-145；
- [7]王亚涛,董兰芳,倪奎.基于 Biharmonic 样条插值的图像渐变算法及实现,中国图像图形学报,2007,12(12):2189-2194；
- [8]黄勇、郭庆荣、任海等,城市土壤重金属污染研究综述[J].热带地点,2005,25（1）：14-18；
- [9]李海滨、林忠辉、刘苏峡，Kriging 法在区域土壤水分估值中的应用.地理研究,2001,20(4):446-452；
- [10]尹骏，上海城郊梯度土壤重金属空间分布特征和评价研究，上海师范大学硕士学位论文，2010.09；
- [11]中国绿色食品发展中心，绿色食品产地环境质量现状评价纲要(试行)，1994；
- [12]黄卫，用 EXCEL 编程计算 Spearman 秩相关系数，环境监测管理与技术，第 13 卷第 3 期，2001.6.；
- [13]尹骏，上海市城郊梯度土壤重金属空间分布特征和评价研究，上海师范大学，硕士学位论文：4；
- [14]MunehD.1993.ConcentrationProfilesofarsenic, cadmium, chromium, copper, lead, mercury, nickel, zinc, vanadium and polynuclear aromatic hydrocarbons(PAH)in forests soil beside an uthan road[J].Sci Total Environ, 218:161-174；
- [15]Dilek G T, Barry J M.2003.Heavy metal contamination in highway soils, comparison of Corpus Christi, Texas and Cincinnati, Ohio shows organic matter is key to mobility[J].Clean Techn Enriron Policy, (4):335-245；
- [16]田秀红，我国城郊蔬菜重金属污染研究进展[J]，食品科学，2009，30(21)：449-453。



## 附录

附录部分无非是程序，文章本身放不下的图片，表格等。

有时间一定要花 30 分钟整理下。显得层次分明，让文章结果的呈现有据可依。~~~

### 附录一：

问题 1 中用于绘制重金属空间分布图像的 Biharmonic(v4) 差值算法的 Matlab 程序代码：

#### Draw\_3D.m

```
%载入地形和浓度数据
x=F1(:,2);
y=F1(:,3);
z=F1(:,4);
z2=F2(:,9);

%产生线性行向量
xi=linspace(1,30000,700);
yi=linspace(1,20000,700);

%利用Biharmonic(v4)算法插值数据
[Xi,Yi]=meshgrid(xi,yi);
Z=griddata(x,y,z,Xi,Yi,'v4');
Z2=griddata(x,y,z2,Xi,Yi,'v4');

%绘制平面浓度分布图
subplot(1,2,1);
pcolor(Xi,Yi,Z2);
shading interp;
set(gca,'DataAspectRatio',[1 1 1]);

%绘制曲面地形图，以浓度分布情况着色
subplot(1,2,2);
surf(Xi,Yi,Z,Z2);
shading interp; colorbar;
set(gca,'DataAspectRatio',[1 1 0.05]);
```

问题 2 中计算重金属之间 Spearman 秩相关系数的 Matlab 程序代码：

#### Spearman.m:

```
function [ R ] = Spearman( x,y )
%SPEARMAN
```

```
R=0;
N = length(x);

%%按升序排列，返回索引值
[x1,i] = sort(x);
[y1,j] = sort(y);

%计算秩系数
D = sum((i-j).^2);
R = 1 - 6*D/(N^3-N);

end

Cal_spearman.m:

load Data
load Means

Ranks_E=zeros(8,8);
Ranks_C=zeros(5,5);

index = ceil(319*rand(30,1));

for i=1:8
    for j=i:8
        Ranks_E(i,j) = Spearman(F2(index,i+1),F2(index,j+1));
    end
end

for i=1:5
    for j=i:5
        Ranks_C(i,j) = Spearman(Means(i,:),Means(j,:));
    end
end

Ranks_E
Ranks_C

clear i j temp
```

问题三中 K 均值聚类模型算法的实现和绘图的 Matlab 程序代码：



```
Draw_K_means.m

clear all;
close all;

% k: 聚类数; eleno: 元素列编号
k=23;
eleno=8;

% 下面绘制浓度插值平面图
% 载入已插值的数据
load (['_griddata/Ele_',int2str(eleno),'_griddata']);

% 创建图形
figure;
hold on;
pcolor(Xi,Yi,Z);
shading interp;
set(gca,'DataAspectRatio',[1 1 1]);
colorbar;

% 清理临时变量
clear Z Xi Yi;

% 下面绘制K-均值计算结果
% 载入数据
load eledatak;
Z=F2(:,eleno+1);
fprintf('Pause for 2 Seconds...\n'); % 暂停两秒
pause(2);

% idx: 数据点聚类索引; c-聚类中心
[idx,c]=kmeans([F1(:,2),F1(:,3),Z],k);

% 随机产生颜色
colr = zeros(k,3);
for i = 1:k
    colr(i,:) = rand(1,3);
end;

for i = 1:k,
    % 绘制所有数据点
    pos = find(idx == i);
```

```

handle = plot3(F1(pos,2), F1(pos,3),F2(pos,eleno + 1),'.');
set(handle,'Color',colr(i,:), 'MarkerSize',3);

% 绘制每个聚类中浓度最高的数据点
posmax = find(F2(pos,eleno + 1) == max(F2(pos,eleno + 1)));
handle2 = plot3(F1(pos(posmax(1)),2),
F1(pos(posmax(1)),3),F2(pos(posmax(1)),eleno), 'x');
set(handle2,'Color',colr(i,:), 'MarkerSize',20);

end;

%设置标题，坐标轴比例
title(['k=',int2str(k)]);
set(gca,'DataAspectRatio',[1 1 0.002]);

%清理临时变量
clear i handle handle2 posmax pos eleno colr;

```

## Calc\_FEM\_Cluster.m

```

%对8类金属进行模糊聚类分析
load data1
r1 = merge(r);
r2 = merge(r1);
r3 = merge(r2);
bh = zeros(8);
bh(find(r2>0.998)) = 1;
bh

function rhat = merge( r )
n=length(r);
for i=1:n
    for j=1:n
        rhat(i,j)=max(min([r(i,:);r(:,j)']));
    end
end
end
end

```