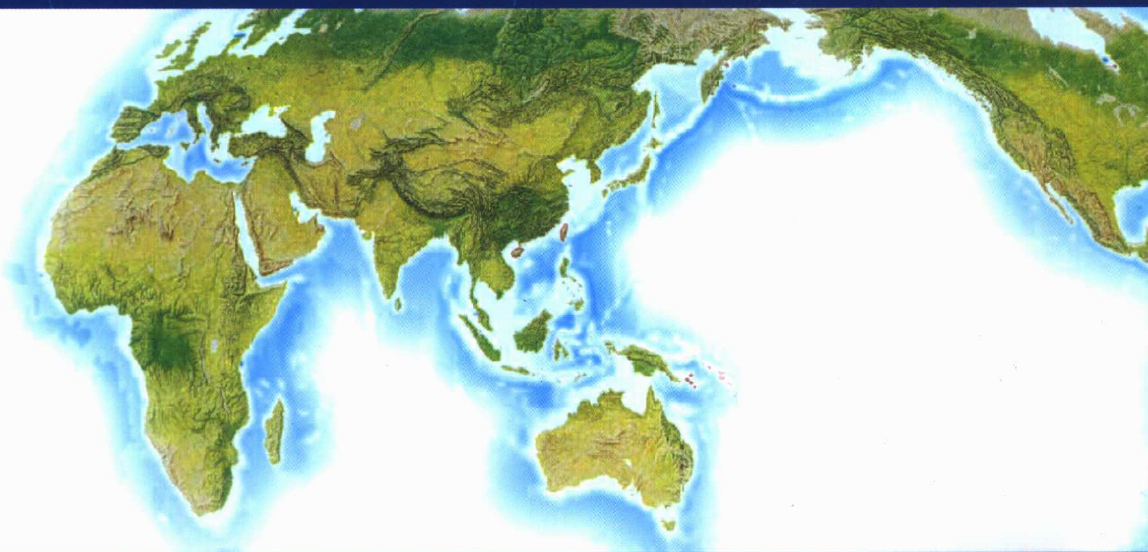
GEOGRAPHIC INFORMATION SCIENCE

高等学校地图学与地理信息系统专业教材

地图数据处理模型的原理与方法

Elements and Methods of Model for Cartographical Data Processing

何宗宜 编著



全国优秀出版社
武汉大学出版社

高等学校地图学与地理信息系统专业教材

地图数据处理模型的原理与方法

Elements and Methods of Model for Cartographical Data Processing

何宗宜 编著

全国优秀出版社
武汉大学出版社

图书在版编目(CIP)数据

地图数据处理模型的原理与方法/何宗宜编著.—武汉：武汉大学出版社,2004.2

高等学校地图学与地理信息系统专业教材

ISBN 7-307-04110-3

Ⅰ.地… Ⅱ.何… Ⅲ.数字模型—应用—地图制图自动化—数据处理 Ⅳ.P282

中国版本图书馆 CIP 数据核字(2003)第 124885 号

责任编辑：解云琳　　责任校对：程小宜　　版式设计：支　笛

出版发行：**武汉大学出版社**　(430072　武昌　珞珈山)

(电子邮件：wdp4@whu.edu.cn　网址：www.wdp.whu.edu.cn)

印刷：武汉理工大印刷厂

开本：787×1092　1/16　　印张：15.75　　字数：368 千字

版次：2004 年 2 月第 1 版　　2004 年 2 月第 1 次印刷

ISBN 7-307-04110-3/P·68　　定价：25.00 元

前　言

数字地图是国家空间数据基础设施的基础。一幅地图表达的地理空间有时是一个地段,有时是一个地区,有时是整个地球,并且具有可量测性、直观性和一览性。另一方面,地理信息系统(GIS)以其丰富的地理信息内容作为数字地图制图生产的基础之一,须满足输出不同比例尺的地图产品的要求,因此需要地图制图综合模型理论和方法来处理地理空间数据。多尺度、多类型、多时态的地理信息是人类研究和解决资源与环境等重大问题时所必需的重要信息资源。随着地理信息系统在社会各个领域的广泛应用,对多种尺度空间数据分析和显示的需求逐渐增加,各部门、机构和单位为解决不同的问题对空间图形数据需求的详细程度是不一样的,GIS 应提供给用户多尺度的空间数据,以提高管理、规划、监测和决策的效率和水平。因此,需要空间数据处理与表示的模型理论和方法知识,使之从单一的较大比例尺派生出较小比例尺或较概略程度的多种比例尺空间数据集。空间分析可以对空间数据进行深加工,向用户提供他们所需要的结果,充分发挥地理信息系统在国民经济建设和国防建设中的作用,所以,空间分析的模型方法是地理信息系统应用的理论基础。随着数字地图制图技术和地理信息系统的快速发展,地图数据处理的理论和方法显得越来越重要。

本书主要介绍了地理信息综合、空间分析和空间数据可视化等地图数据处理模型的原理和方法,是作者 20 余年科研和教学成果的积累。学生通过学习可基本掌握地图数据处理的理论和方法,为今后在实际工作中的数字地图和地理信息系统设计与应用打下坚实基础。

书中插图由何晶、张琳、陶利佳、白菁、谭芬、关焱、赵娟、赵嵘、曹钦、刘祥、唐云妹等绘制。书中还引用了许多参考资料,在参考文献中未一一列出,在此一并致谢!

由于作者水平所限,书中疏漏之处敬请读者批评指正。

何宗宜

2004 年 2 月于珞珈山

目　　录

第一章　概　　述

现代地图制图以空间数据为主要对象,数据处理是当前地图制图的主要研究领域。地图制图数据处理模型的原理与方法是地图制图与数学模型相结合的一门边缘学科,它应用数学方法处理制图数据,用相应的地图制图模型表达数据处理的结果。

§1-1　地图制图数据处理模型的发展

数学方法在地图制图中的应用有着很长的历史。公元前3世纪地图上就使用了较严密的数学方法,开始出现经纬线。由于多方面的原因,两千多年来,数学方法在地图制图中的应用仅局限于地图的数学基础方面,使得地图制图学的大部分领域长期处于定性研究阶段,因而被人们称为"定性科学"或"经验科学"。

20世纪40年代,数学方法在地图制图学中的应用开始出现了良好的转机,利用图解计算法和数理统计方法研究地图要素的选取获得了较好的效果。前苏联地图制图学家根据地图载负量和视觉变量首先提出图解计算法,较好地解决了居民地选取指标的确定,接着较系统地将数理统计引入地图制图。保查罗夫(M. K. Ъоцалов)等人于1957年发表了《制图作业中的数理统计方法》专著,运用数理统计方法研究了地理要素的分布规律和这些要素的制图综合选取指标。60年代,德国的特普弗尔(F. Töpfer)多次发表论文,主张用资料地图和新编地图比例尺分母之比的开方根作为确定新编地图地物选取指标的数学依据,提出了选取规律公式;70年代发表了《制图综合》专著,将开方根规律选取公式系统化。1968年,捷克斯洛伐克的斯恩卡(E. Srnka)用相关解析法,建立了顾及地物大小和地物密度变化的选取公式。但是,这些研究只是解决了选取数量问题,我们称这类制图综合数学模型为"定额选取数学模型"。1976年,前苏联地图制图学家鲍罗金(A. B. Ъоролин)用地图物体本身的大小和所处的地理环境(物体密度)两个标志来衡量地图物体的重要性,确定地图物体的取舍。该方法还可以确定具体选取哪一个物体,使地图制图综合模型大大地进了一步,我们称这类制图综合数学模型为"结构选取数学模型"。

在我国,20世纪50年代末和60年代初,不少地图制图学者用图解计算法和数理统计法研究居民地、河流、道路网等选取指标的数学模型,取得了一些成果。70年代以来,有人着手用回归分析方法研究确定居民地和河流选取的数学模型,后来又有些学者利用多元回归分析方法建立地图物体的制图综合数学模型,使确定制图地物选取指标模型的精度有了很大提高。从80年代中期开始,有人利用模糊数学和图论来研究居民地、河流和道路网等的"结构选取数学模型",也取得了不少成果。

地图制图数据处理模型不仅限于地图制图综合数学模型的研究,实际上数学方法在专题地图制图数据处理中应用也比较广泛。从20世纪50年代以来,地学研究方法发生变革,

从定性分析发展到定量分析。60 年代，有人将多元统计分析应用于地学领域，推动了地图制图数据处理模型的发展。70 年代，不少地图制图学者应用统计分析和信息论分析地图内容，在此基础上，形成了比较系统的地图制图数据处理模型的理论与方法，针对不同的问题，提出了相应的数学模型和地图制图相结合的途径。80 年代，有些学者开始把模糊数学、最优化方法等现代数学引入专题地图制图的研究领域，取得了不少研究成果。

从 20 世纪 90 年代以来，许多地图制图学者利用数学形态学、分形理论和小波理论等现代数学对空间数据多尺度处理与表示进行深入地探讨，取得了许多研究成果，使地图制图数据处理模型得到进一步发展。

§1-2 地图制图数据处理模型的应用

地图制图数据处理模型的研究虽然还处在初级阶段，但随着它的逐步发展，已在地图制图综合、地图设计、专题地图制图和空间数据处理中发挥着越来越重要的作用。

地图制图综合是地图制图学重要的研究课题之一，是一种特殊的地图制图数据处理方法。随着地理信息技术的发展，地图制图综合的数学模型已显示出广阔的应用前景。地图制图综合模型主要有定额选取模型、结构选取模型和图形化简模型。

1. 定额选取模型

普通地图，特别是国家基本比例尺地形图的编制，都是利用大比例尺实测地形图逐步缩编而成，用数学模型确定缩小后的新编图的地物选取数量，可提高地图制图综合的质量和科学性。定额选取模型主要有图解计算法、方根模型和数理统计模型。

2. 结构选取模型

该模型是确定选取具体地图制图物体的数学模型。根据制图物体的结构关系，从大比例尺资料图上的制图物体中寻找出更重要的一部分物体表示在新编地图上。从地物的层次关系(等级关系)、空间关系(毗邻与包含)和拓扑关系(邻接和关联)等方面来解决具体选取哪些物体的问题。结构选取数学模型主要有等比数列法、模糊数学模型和图论模型。

3. 图形化简模型

该模型是对已选取的制图物体的平面图形进行化简，并保持平面图形的主要形状特征的数学模型。

地图制图数据处理模型在专题地图制图中的应用主要有地图制图要素的分级模型、地图制图要素的相关模型、地图制图要素空间分布趋势模型、地图制图要素预测模型、地图制图要素信息简化模型和地图制图要素类型划分模型。

1. 地图制图要素的分级模型

地图制图要素的分级模型主要解决分级数的确定和分级界线的确定两个方面的问题。其中分级界线的确定是分级的核心问题，它对分级表示能否保持数据特征起决定性作用。对制图要素的数量特征进行科学的分级，有利于研究要素空间分布的趋势和差异。针对不同的目的要求和数据分布本身的特征，采用不同的数学模型。

2. 地图制图要素的相关模型

地图制图要素的相关模型研究要素之间统计相关关系，并编制显示要素或现象(地区或部门)相互关系(结构特征)的相关地图。

3. 地图制图要素空间分布趋势模型

地图制图要素空间分布趋势模型研究要素空间分布的规律和特点，并编制要素空间分布趋势图（背景面图和剩余面图）。

4. 地图制图要素预测模型

地图制图要素预测模型用回归分析方法研究现象间的制约关系，确定一个变量随其他变量的变化而变化的规律，编制预测地图。

5. 地图制图要素信息简化模型

地图制图要素信息简化模型应用主因素分析和主成分分析，在地图上的多维信息中区分出最为实质的部分，寻找出最重要的地理规律。信息简化的数学模型可以简化变量结构，编制综合地图。

6. 地图制图要素类型划分模型

地图制图要素类型划分模型用聚类分析方法研究样品或变量的组合分类，编制类型图、区划图、分类图。

当前 GIS 数据库为了满足人们应用空间数据集的不同需求，不得不存储多种来源、多种比例尺、多种详细程度的空间数据，造成多重表示现象，从而会产生大量数据冗余以及内存开销的增加等相关的弊端，更重要的是在进行跨比例尺综合分析时会产生严重的数据矛盾。因此，需要研究空间数据多尺度处理与表示的模型方法，使之在从单一的较大比例尺或较详细程度的空间数据集派生较小比例尺或较概略程度的多种比例尺空间数据集时，通过多尺度变换，能够从一种表示完备地过渡到另一种表示，这种完备性的要求就是派生过程要保持相应尺度的空间精度和空间特征，保证空间关系不发生变化。空间数据多尺度处理数学形态学模型、分形理论模型和小波理论模型运用现代数学方法，部分解决了空间图形数据的多尺度处理与表示。

另外，地图要素分布规律的数学模型，用于对地图要素分布规律进行模拟；地图制图评价数学模型，用于对地图资料、地图制图综合程度、地图制图分类分级的评价。这些模型为地图设计、地图制图综合和地图制图数据处理提供了科学依据。

第二章 地图制图数据处理模型的数学基础

地图制图数据处理模型是以数学作为基础的。由于地图制图数据处理极为复杂，它涉及许多数学问题，比较集中的有数理统计、模糊数学、信息论、图论、数学形态学、分形理论和小波理论。这里仅介绍与地图制图数据处理有关的数学基本概念，更深入的数学内容请参考有关数学专著。数学形态学、分形理论和小波理论的数学基础将在相关章节进行介绍。

§2-1 地图制图数据处理模型的数理统计基础

数理统计是地图制图数据处理模型的重要数学基础，是研究地图制图综合，制图要素的分布规律、相互关系、组合特征和发展变化趋势的重要数学工具。

一、地图制图数据类型

地图制图要素的特征和性质是用数据来表示的。地图制图数据可分为定量数据和定性数据。

1. 定量数据

定量数据是一种连续量，可分为间隔尺度数据和比率尺度数据。

(1)间隔尺度数据

间隔尺度数据是有实际单位的度量，如米表示长度，千克表示质量等。间隔尺度数据是地图制图中常见的一种数据类型，一般数学方法和数理统计都以这类数据为基础。

(2)比率尺度数据

比率尺度数据是以一个基准量作为衡量标准的数据，如百分比、百分含量、某种比值。比率尺度数据也是地图制图中常见的一种数据类型，可用于一般数学方法和数理统计处理。

2. 定性数据

定性数据是一种不连续量，可分为有序尺度数据、二元数据和名义尺度数据。

(1)有序尺度数据

有序尺度数据只表示制图物体的次序和等级关系，不表示具体的数量。如居民地的行政等级有首都、省会、地级市、县、乡(镇)、村，虽然没有表示出具体的数据，但却可以分出行政等级的高低和次序。

(2)二元数据

二元数据是表示地图上的图斑是否具有某种特性。如研究地图上的某个图斑所代表的区域在相应实地上是否具有某种树木，可以用数据“0”表示没有，“1”表示有。通过二元数据，可以把地图制图的定性数据和定量数据联系起来进行数量分析。

(3)名义尺度数据

名义尺度数据是表示制图物体的性质差异的数据，例如土壤分类、土地利用分类、植被分类等。

二、地图制图数据的数字特征

1. 频数与频率

频数与频率表示制图要素分布基本特征。设有一组数据为 $x_1, x_2, \cdots, x_n$，按一定的间距分组。在各组出现的次数称为频数，用 f_i 表示；各组频数与总频数之比叫频率，用 p_i 表示。计算公式如下：

$$p_i = f_i \Big/ \sum_{i=1}^{n} f_i \tag{2-1}$$

2. 平均数、数学期望、中位数和众数

这些数据是表示地图制图数据分布的集中趋势。

(1)平均数

平均数是表示地图制图数据分布的集中位置，用 $\bar{x}$ 表示。设有一组数据为 $x_1, x_2, \cdots, x_n$，则 $\bar{x}$ 为：

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2-2}$$

(2)数学期望

数学期望 M_x 是以概率 p_i 为权的加权平均数，表示为：

$$M_x = \sum_{i=1}^{n} p_i x_i \tag{2-3}$$

(3)中位数

中位数是按数值大小排列的中间数，偶数列则为中间两个数的平均值。

(4)众数

众数是出现次数最多的某一数值。

3. 极差、离差、方差和变异系数

这些数值是反映数据的离散程度。

(1)极差

极差是最大值与最小值的差值。

(2)离差

离差是各数值与平均值之差，用公式表示为：

$$d_i = x_i - \bar{x} \tag{2-4}$$

离差绝对值的平均值称为平均离差，用公式表示为：

$$M_d = \sum_{i=1}^{n} \left| x_i - \bar{x} \right| \Big/ n \tag{2-5}$$

离差平方和为：

$$d^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{2-6}$$

(3)方差

方差是用离差平方和除以样本容量得出的，它是反映各数值与平均值的离散程度的重要指标，用公式表示为：

$$\sigma^2 = \sum_{i=1}^{n}(x_i - \bar{x})^2 \Big/ n \tag{2-7}$$

标准差是方差的平方根，当用样本标准差对总体标准差进行估计时，则采用无偏估计值，即：

$$S = \sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \Big/ (n-1)} \tag{2-8}$$

(4)变异系数

变异系数是衡量要素的相对变化(波动)的程度。即：

$$C_v = \frac{S}{\bar{x}} \times 100\% \tag{2-9}$$

三、地图制图数据的分布特征参数

地图制图数据处理中常用偏态系数和峰态系数来衡量制图数据的分布特征。

1. 偏态系数

偏态系数表示要素分布的不对称性。偏态系数的计算公式为：

$$C_v = \frac{\mu_3}{S^3} \tag{2-10}$$

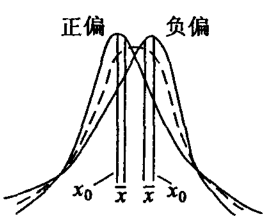


图 2-1 偏态

式中，μ_3 是三阶中心矩，即：

$$\mu_3 = \sum_{i=1}^{n}(x_i - \bar{x})^3 \Big/ n$$

S 为标准差。当 $C_v > 0$ 时，众数在平均值的左边，称为正偏；$C_v < 0$时，众数在平均值的右边，称为负偏；当 $C_v = 0$ 时，图形对称(如图 2-1)。

2. 峰态系数

峰态系数是表示分布图形的峰度高低，是要素分布在均值附近的集中程度。峰态系数的计算公式为：

$$C_e = \frac{\mu_4}{S^4} \tag{2-11}$$

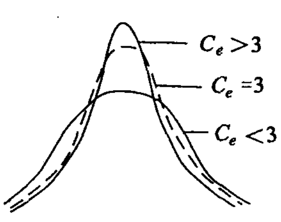


图 2-2 峰态

式中，μ_4 是四阶中心矩，即：

$$\mu_4 = \sum_{i=1}^{n}(x_i - \bar{x})^4 \Big/ n$$

S 为标准差。标准正态分布时，$C_e = 3$；当 $C_e > 3$ 时，称为高峰态；$C_e < 3$ 时，称为底峰态(如图 2-2)。

四、地图制图数据的标准化

地图制图数据源的每个样本有多种变量的原始数据，各种变量的量纲和数量大小是很不一致的，变化幅度也不一样。如直接用原始数据进行计算，就

会突出绝对值大的变量的作用。为了给每种变量以同一量度,使地图制图数据处理结果能客观地反映实际,建立模型前需要对原始数据进行标准化。

1. 标准差标准化

设有 n 个单元,每个单元有 m 个数据,每个变量可记为 x_{ij};$i=1,2,\cdots,n$;$j=1,2,\cdots,m$。标准化后的变量 x'_{ij} 为:

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j} \tag{2-12}$$

式中,$\bar{x}_j$ 为第 j 个变量的平均数;S_j 为第 j 个变量的标准差。经过标准差标准化后,每种变量平均值为 0,标准差为 1。

2. 极差标准化

标准差标准化要计算标准差,为了方便,也可采用极差标准化,把变量变换到 0~1 范围之内。标准化后的变量 x'_{ij} 为:

$$x'_{ij} = \frac{x_{ij} - x_{j\min}}{x_{j\max} - x_{j\min}} \tag{2-13}$$

式中,$x_{j\max}$ 和 $x_{j\min}$ 分别是第 j 个变量的最大值和最小值。

五、回归分析

回归分析是地图要素分布规律的数学模型、地图制图综合定额选取数学模型、地图制图要素空间分布趋势数学模型和地图制图要素动态分析和预测数学模型等的数学基础。

1. 一元线性回归

(1)回归方程

一元线性回归主要是处理两个制图变量 x 与 y 之间的线性关系。通过观测或试验可得到两个变量 x 与 y 的若干数据 x_i 与 $y_i(i=1,2,\cdots,n)$,并将它们绘到坐标纸上,可以看出其分布近似一条直线(如图 2-3)。设这条直线方程为:

$$\hat{y} = a + bx \tag{2-14}$$

$\hat{y}$ 式中,a,b 是待定的系数。对于每个 x_i 所对应的观测值 y_i 与

$$\hat{y}_i = a + bx_i$$

之间有误差

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

若有一条直线能使所有误差平方和达到最小,那么,这条直线就称为回归直线。

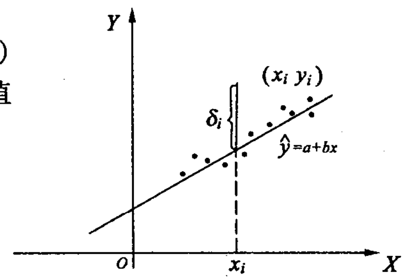


图 2-3

令误差平方和为 Q,则

$$Q = \sum_{i=1}^{n} \delta_i^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - a - bx_i)^2 \tag{2-15}$$

要使 Q 为最小,由微积分中求极值的原理可知,只要将 Q 分别对 a,b 求偏导数,然后令偏导数为 0,即可求出 a,b,令

$$\frac{\partial Q}{\partial a} = -2\sum_{i=1}^{n} (y_i - a - bx_i) = 0$$

$$\frac{\partial Q}{\partial b} = -2\sum_{i=1}^{n}(y_i - a - bx_i)x_i = 0$$

即

$$\sum_{i=1}^{n}(y_i - a - bx_i) = 0 \tag{2-16}$$

$$\sum_{i=1}^{n}(y_i - a - bx_i)x_i = 0 \tag{2-17}$$

由(2-16)式解出

$$a = \bar{y} - b\bar{x} \tag{2-18}$$

其中

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n}x_i, \qquad \bar{y} = \frac{1}{n}\sum_{i=1}^{n}y_i$$

将(2-18)式代入(2-17)式,得

$$b = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{L_{xy}}{L_{xx}} \tag{2-19}$$

把(2-19)式代入(2-18)式,就可求出 a。求出 a,b 后,便可写出 x 与 y 的关系式:

$$\hat{y} = a + bx$$

上式就称为 y 对 x 的回归方程。

(2)回归方程的显著性检验

对于任何一组观测数据(x_i,y_i),均可按上述方法建立回归方程。如果变量 x 与 y 之间的关系是线性的,回归方程才有意义;若变量 x 与 y 不是线性关系,回归方程毫无意义。因此,需要对回归方程进行相关显著性检验。对一元线性回归方程一般采用相关系数检验法。相关系数

$$r = \frac{\sum_{i=1}^{n}x_iy_i - \frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i}{\sqrt{\left[\sum_{i=1}^{n}x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2\right]\left[\sum_{i=1}^{n}y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n}y_i\right)^2\right]}} \tag{2-20}$$

若计算出的 r 值大于查表(相关系数检验表)值 r_α,则认为相关显著,回归方程有意义。α 是置信水平,$1-\alpha$ 代表置信度,如 $\alpha=0.05$,表示可靠程度为95%。

2. 一元非线性回归

(1)回归方程

实际地图制图数据处理中,变量之间的关系常常是非线性的,如地图上居民地选取指标与居民地密度之间就是幂函数关系。一元非线性回归方程的求法一般是通过数学变换,使非线性关系化为线性关系,再利用线性回归的方法解非线性回归方程。

当变量 x 与 y 之间是幂函数关系(如图2-4)

$$y = ax^b$$

对上式两边取对数,得

$$\ln y = \ln a + b\ln x$$

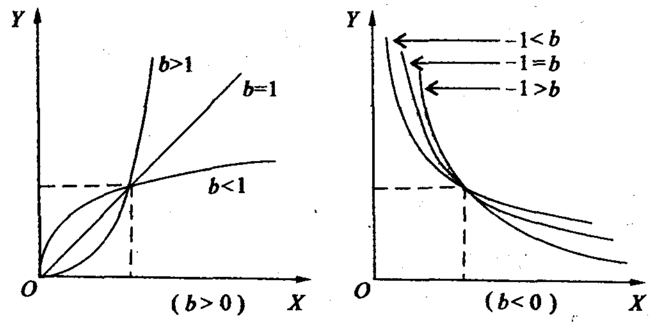


图 2-4

令

$$\ln y = Y,\ \ln a = A,\ \ln x = X$$

于是有

$$Y = A + bX$$

根据(2-18)式、(2-19)式有

$$\left.\begin{aligned} b &= \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n}(X_i - \overline{X})^2} \\ A &= \overline{Y} - b\overline{X} \\ a &= e^A \end{aligned}\right\} \tag{2-21}$$

(2)回归方程的显著性检验

一元非线性回归方程的显著性检验,是检验变量 x 与 y 之间的关系是否有非线性回归方程建立的这种非线性关系,即观测点 (x_i, y_i) 与曲线的拟合程度的好坏,一般用相关指数进行检验。相关指数

$$R = \sqrt{1 - \frac{Q}{L_{yy}}} \tag{2-22}$$

式中

$$Q = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

$$L_{yy} = \sum_{i=1}^{n}(y_i - \overline{y})^2$$

若计算出的 R 值大于查表值 R_α,则认为相关显著,回归方程有意义。

3. 多元线性回归分析

在地图制图数据处理中,影响制图物体的因素往往有多种,需要进行多元线性回归分析。多元线性回归分析的基本原理与一元线性回归分析的原理相同。

(1)多元线性回归方程

设有 p 个自变量 $x_1, x_2, \cdots, x_p$ 与因变量 y,它们有如下关系式:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \cdots + b_px_p \tag{2-23}$$

式中，$b_0,b_1,\cdots,b_p$ 为待定参数。

设对变量 $x_1,x_2,\cdots,x_p,y$ 作了 n 次观测，其中第 k 次观测数据为：

$$\hat{y}_k = b_0 + b_1x_{k1} + b_2x_{k2} + \cdots + b_px_{kp}$$

令

$$Q = \sum_{k=1}^{n}(y_k - \hat{y}_k)^2 = \sum_{k=1}^{n}(y_k - b_0 - b_1x_{k1} - b_2x_{k2} - \cdots - b_px_{kp})^2$$

将 Q 分别对 $b_0,b_1,\cdots,b_p$ 求偏导数，得 $p+1$ 个方程

$$\frac{\partial Q}{\partial b_0} = 2\sum_{k=1}^{n}(y_k - b_0 - b_1x_{k1} - b_2x_{k2} - \cdots - b_px_{kp})(-1) = 0$$

$$\frac{\partial Q}{\partial b_1} = 2\sum_{k=1}^{n}(y_k - b_0 - b_1x_{k1} - b_2x_{k2} - \cdots - b_px_{kp})(-x_{k1}) = 0$$

$$\vdots$$

$$\frac{\partial Q}{\partial b_j} = 2\sum_{k=1}^{n}(y_k - b_0 - b_1x_{k1} - b_2x_{k2} - \cdots - b_px_{kp})(-x_{kj}) = 0$$

$$\vdots$$

$$\frac{\partial Q}{\partial b_p} = 2\sum_{k=1}^{n}(y_k - b_0 - b_1x_{k1} - b_2x_{k2} - \cdots - b_px_{kp})(-x_{kp}) = 0$$

由第一个方程得

$$b_0 = \bar{y} - b_1\bar{x}_1 - \cdots - b_p\bar{x}_p \tag{2-24}$$

式中

$$\bar{y} = \frac{1}{n}\sum_{k=1}^{n}y_k, \quad \bar{x}_j = \frac{1}{n}\sum_{k=1}^{n}x_{kj}, \quad j = 1,2,\cdots,p, \quad k = 1,2,\cdots,n$$

将 b_0 代入后面 p 个方程，整理得

$$\left.\begin{array}{l} L_{11}b_1 + L_{12}b_2 + \cdots + L_{1p}b_p = L_{1y} \\ L_{21}b_1 + L_{22}b_2 + \cdots + L_{2p}b_p = L_{2y} \\ \vdots \\ L_{p1}b_1 + L_{p2}b_2 + \cdots + L_{pp}b_p = L_{py} \end{array}\right\} \tag{2-25}$$

式中

$$L_{ii} = \sum_{k=1}^{n}(x_{ki} - \bar{x}_i)^2, \quad i = 1,2,\cdots,p$$

$$L_{ij} = \sum_{k=1}^{n}(x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j), \quad i,j = 1,2,\cdots,p$$

$$L_{iy} = \sum_{k=1}^{n}(x_{ki} - \bar{x}_i)(y_k - \bar{y}), \quad i = 1,2,\cdots,p$$

由(2-25)式解出 $b_1,b_2,\cdots,b_p$，根据(2-24)式可得 b_0。

如果只有 y,x_1,x_2 三个变量，则(2-25)式可变为：

$$\left.\begin{array}{l} L_{11}b_1 + L_{12}b_2 = L_{1y} \\ L_{21}b_1 + L_{22}b_2 = L_{2y} \end{array}\right\} \tag{2-26}$$

由(2-26)式得