

■ 编著 樊民强

解方程得  
所以

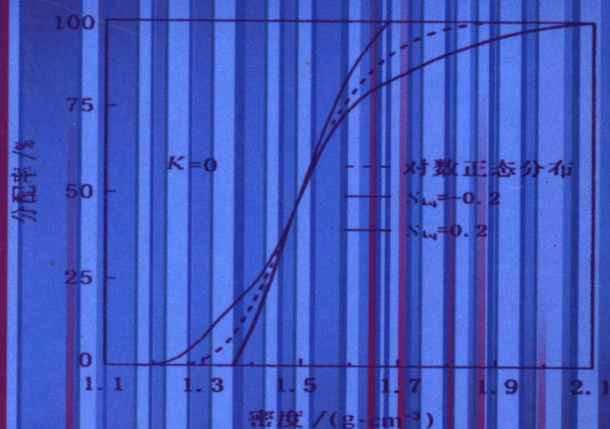
$$\frac{0.6745}{D_q}(x_{25} - x_{50} - \Delta) = -0.6745$$

$$\frac{0.6745}{D_q}(x_{75} - x_{50} - \Delta) = 0.6745$$

$$\Delta = \frac{x_{75} + x_{25} - 2x_{50}}{2}$$

$$u = \frac{0.6745}{D_q}(x - x_{50} - \Delta) =$$

# 选煤数学模型与数据处理



煤炭工业出版社

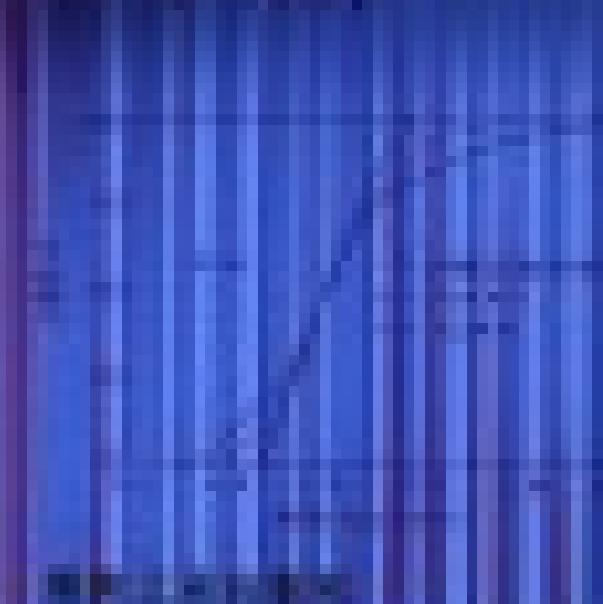


中国矿业大学  
北京

# 选煤数学模型与数据处理

王德明 王德成 王德成 王德成

中国矿业大学  
北京



中国矿业大学出版社

# 选煤数学模型与数据处理

樊民强 编著

煤炭工业出版社

· 北 京 ·

**图书在版编目 (CIP) 数据**

选煤数学模型与数据处理/樊民强编著. —北京: 煤炭工业出版社, 2004

ISBN 7-5020-2552-9

I. 选… II. 樊… III. ①选煤—数学模型②选煤数值计算 IV. TD941

中国版本图书馆CIP数据核字 (2004) 第107330号

煤炭工业出版社 出版  
(北京市朝阳区芍药居35号 100029)  
网址: [www.cciph.com.cn](http://www.cciph.com.cn)  
煤炭工业出版社印刷厂 印刷  
新华书店北京发行所 发行

\*

开本 850mm×1168mm  $1/32$  印张  $8\frac{1}{2}$   
字数 214千字 印数 1—1,200  
2005年4月第1版 2005年4月第1次印刷  
社内编号 5323 定价 30.00元

**版权所有 违者必究**

本书如有缺页、倒页、脱页等质量问题, 本社负责调换

## 目 录

|                       |    |
|-----------------------|----|
| <b>1 数学模型概论</b>       | 1  |
| 1.1 数学模型的定义           | 1  |
| 1.2 数学模型的分类           | 3  |
| 1.2.1 根据模型的来源分类       | 3  |
| 1.2.2 根据模型中变量和时间的关系分类 | 4  |
| 1.2.3 根据模型中变量的性质分类    | 4  |
| 1.3 建立数学模型的方法和步骤      | 5  |
| 1.3.1 建立数学模型的方法和一般步骤  | 5  |
| 1.3.2 建立经验模型的一般方法     | 6  |
| <b>2 数理统计基础</b>       | 9  |
| 2.1 随机变量与分布           | 9  |
| 2.1.1 随机过程与随机变量       | 9  |
| 2.1.2 随机变量的数字特征       | 10 |
| 2.1.3 正态分布与相关分布       | 13 |
| 2.1.4 其它连续型分布         | 20 |
| 2.2 统计分析              | 28 |
| 2.2.1 总体与样本           | 28 |
| 2.2.2 区间估计            | 29 |
| 2.2.3 假设检验            | 33 |
| <b>3 回归分析</b>         | 40 |
| 3.1 引言                | 40 |
| 3.2 一元线性回归模型          | 40 |
| 3.2.1 回归系数的确定         | 41 |
| 3.2.2 模型显著性检验         | 42 |

|                              |            |
|------------------------------|------------|
| 3.2.3 回归方程的精度 .....          | 44         |
| 3.3 可线性化曲线的线性回归 .....        | 45         |
| 3.3.1 可线性化曲线的基本类型与线性转化 ..... | 46         |
| 3.3.2 模型的检验 .....            | 47         |
| 3.3.3 程序 .....               | 50         |
| 3.4 线性方程组的解法 .....           | 55         |
| 3.4.1 高斯消元法解线性方程 .....       | 55         |
| 3.4.2 高斯消元法求逆矩阵 .....        | 59         |
| 3.5 一元多项式回归分析 .....          | 62         |
| 3.5.1 多项式阶数的判断 .....         | 62         |
| 3.5.2 模型参数的估计 .....          | 64         |
| 3.5.3 曲线平滑程序 .....           | 64         |
| 3.6 多元线性回归分析 .....           | 70         |
| 3.6.1 多元线性回归 .....           | 70         |
| 3.6.2 回归方程的显著性检验 .....       | 72         |
| 3.6.3 回归系数的检验 .....          | 74         |
| 3.7 逐步回归分析 .....             | 74         |
| 3.7.1 逐步回归 .....             | 74         |
| 3.7.2 逐步降元回归 .....           | 75         |
| 3.7.3 逐步增元回归 .....           | 82         |
| 3.8 非线性回归 .....              | 95         |
| 3.8.1 高斯-牛顿法 .....           | 95         |
| 3.8.2 非线性阻尼最小二乘法 .....       | 97         |
| 3.8.3 非线性回归程序与示例 .....       | 97         |
| <b>4 插值法 .....</b>           | <b>105</b> |
| 4.1 拉格朗日插值 .....             | 106        |
| 4.1.1 拉格朗日插值多项式通式 .....      | 106        |
| 4.1.2 拉格朗日插值计算方法 .....       | 108        |
| 4.1.3 用拉格朗日插值加密浮沉资料 .....    | 112        |

---

|       |                               |     |
|-------|-------------------------------|-----|
| 4.2   | 牛顿插值 .....                    | 116 |
| 4.2.1 | 差商与差分 .....                   | 116 |
| 4.2.2 | 牛顿基本插值公式 .....                | 118 |
| 4.3   | 埃尔米特插值 .....                  | 123 |
| 4.4   | 样条插值 .....                    | 129 |
| 5     | 过程优化 .....                    | 139 |
| 5.1   | 非线性方程的数值解 .....               | 139 |
| 5.1.1 | 单变量非线性方程的数值解 .....            | 139 |
| 5.1.2 | 非线性方程组的数值解 .....              | 146 |
| 5.2   | 单变量优化的黄金分割法 .....             | 149 |
| 5.3   | 线性规划 .....                    | 156 |
| 5.3.1 | 线性规划的数学模型 .....               | 156 |
| 5.3.2 | 两变量线性规划的图解法 .....             | 157 |
| 5.3.3 | 线性规划的标准形式 .....               | 158 |
| 6     | 重选数学模型 .....                  | 171 |
| 6.1   | 可选性曲线数学模型及计算机绘图方法 .....       | 171 |
| 6.1.1 | 密度曲线数学模型 .....                | 171 |
| 6.1.2 | 灰分曲线数学模型 .....                | 172 |
| 6.1.3 | 可选性曲线的计算机绘制 .....             | 173 |
| 6.1.4 | 密度曲线特征参数与煤炭可选性关系研究 .....      | 184 |
| 6.2   | 分配曲线的数学模型 .....               | 194 |
| 6.2.1 | 分配率的计算 .....                  | 194 |
| 6.2.2 | 分配曲线的特点 .....                 | 196 |
| 6.2.3 | 分配曲线数学模型及其算法 .....            | 200 |
| 6.2.4 | 以四分位数作模型参数的分配曲线数学模型 .....     | 202 |
| 6.3   | 颗粒在跳汰床层中分布形态及数学模型<br>研究 ..... | 209 |
| 6.3.1 | 试验设计 .....                    | 210 |
| 6.3.2 | 颗粒在床层中基本分布形态的数学描述 .....       | 210 |

|          |                                |            |
|----------|--------------------------------|------------|
| 6.3.3    | 颗粒在床层中分布的理论模型 .....            | 214        |
| 6.3.4    | 四分位数模型对颗粒在床层中分布形态的<br>表征 ..... | 217        |
| 6.3.5    | 颗粒在床层中分布的经验模型——B分布 .....       | 221        |
| <b>7</b> | <b>选煤过程优化计算 .....</b>          | <b>226</b> |
| 7.1      | 重选流程的优化计算 .....                | 226        |
| 7.1.1    | 重选产品指标的基本算法 .....              | 226        |
| 7.1.2    | 重选流程优化程序设计 .....               | 229        |
| 7.2      | 原煤入选方案比较 .....                 | 239        |
| 附录1      | 标准正态分布表 .....                  | 248        |
| 附录2      | $\chi^2$ 分布 .....              | 250        |
| 附录3      | $t$ 分布表 .....                  | 253        |
| 附录4      | $F$ 分布表 .....                  | 255        |
| 附录5      | 相关系数表 $R_a$ .....              | 261        |
| 参考文献     | .....                          | 262        |

# 1 数学模型概论

## 1.1 数学模型的定义

选煤是一个利用若干种分选与分离设备，将原煤分成不同质量产品的工艺过程。具体的工艺一般都会在选煤厂调度室用模板或屏幕显示出来。如果说屏幕上显示的设备联系图是对选煤过程的模型化描述，则设备联系图上的各个设备就是实际设备的模型。模型总是与它描述的现实对象对应存在的。

原型指客观存在的实体、过程等实际对象。模型则指为了某个特定的目的将原型的某一部分信息简缩、提炼而构成的原型替代物。模型和原型之间具有一定的相似性，这种相似性可以是外形相似、内部结构相似，也可以是行为相似。

用模型替代原型的方式来分类，模型可以分成物质模型（形象模型）和理想模型（抽象模型）。前者包括直观模型、物理模型等，后者包括思维模型、符号模型、数学模型等。

直观模型通常是把原型的尺寸按比例缩小或放大，主要追求外观形象上的逼真。如供展览的实物模型、玩具等。

物理模型主要指科技工作者根据与原型相似的原理构造的模型，它不仅可以显示原型的外形或某些特征，而且可以用来进行模型试验，间接地研究原型的某些规律。如风洞试验中的飞机模型，核爆炸反应模型等。利用物理模型通常可得到很有价值的结果，但也存在成本高、时间长、不灵活等缺点。

思维模型指人们通过对原型的反复认识，将获取的知识以经验的形式直接储存于大脑中，并根据思维或直觉作出相应的决策。如汽车司机驾驶汽车，领导凭经验作决策等，就利用了思维模型。

符号模型是在一些约定或假设条件下借助专门的符号、线条等,按一定形式组合起来描述原型,如地图、电路图、工艺流程图等。

数学模型则是由数字、字母或其它数学符号组成的,描述过程或运动物体特征的因果关系的数学表达式、图表或算法。

数学模型融会在人类社会的方方面面。如对未来某个时期人口数变化、能源需求的预测、具体工业过程的控制等,都离不开数学模型。甚至是日常生活中的旅游和购物,人们自觉不自觉地也会利用数学模型确定一个优化方案。

一般地,数学模型可以描述为:对于一个现实对象,为了一个特定目的,根据其内在规律,作出必要的简化假设,运用适当的数学工具,得到的一个数学结构。

例如:分配曲线是不同成分(密度级或粒度级)在某一产品中的分配率的图示,是表示分离效果的特性曲线,其中,不同成分在重产物中的分配曲线是一条极小值为0,极大值为100%的S形增长曲线,类似于概率分布累计曲线。为了简便描述该曲线的特点,人们假设重介分选的分配曲线服从于正态分布,跳汰分选的分配曲线服从于对数正态分布。根据实验数据,绘制分配曲线,从曲线上可以查到一些特殊的点。分配曲线上分配率50%对应的密度,在选煤上称为实际分选密度,记为 $\delta_p$ 。分配曲线上分配率75%和25%对应的密度,可以构成表征实际分选相对于理想分选的偏离程度的指标可能偏差 $E$ ,  $E = (\delta_{75} - \delta_{25}) / 2$ 。由分选密度和可能偏差,就可以构建起重介分配曲线的数学模型

$$E(\delta) = \int_{-\infty}^{\delta} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt, t = \frac{0.6745}{E_p} (\delta - \delta_p) \quad (1.1)$$

在上述问题的描述中,研究的对象是分配曲线,目的是为了从中抽象出特征指标,以表征分选效果和对分选过程进行预测,S形增长是分配曲线的基本规律,分配曲线服从于正态分布是对曲线性质作出的假设,通过图中查找特殊的点可以确定分配曲线的

特性参数, 最终得到分配曲线的正态分布数学模型。

建立数学模型的目的是解决生产过程的计算、预测和优化问题, 为设计、管理和控制提供所需的数据。

## 1.2 数学模型的分类

数学模型可以按照不同的方式分类。下面介绍几种分类方法。

### 1.2.1 根据模型的来源分类

(1) 理论模型。根据实体内部的物理和化学性质, 通过分析推导出来的模型。

适用于作用机理清楚、因果关系明确的生产过程。在选煤中实际应用较少, 一般只涉及化学反应或水力分级作业。

(2) 经验模型。经验模型是指不考虑实际内部的变化(或机理不清楚, 内部变化暂不清楚), 只着重于外部的关系, 把收集到的输入和输出观测值, 用数理统计的方法, 导出输入、输出变量之间的关系, 建立数学模型。它适用于作用机理不清楚、因果关系不明确的生产过程。在选矿中的实际应用广泛, 如粒度方程、可选性曲线模型等。

(3) 综合模型。模型结构来自理论分析, 但其中的某些参数未确定, 需要收集现场生产数据或通过试验, 用数学方法来确定。如在筛分动力学研究中, 根据单位时间内透筛量与筛上物中可透筛颗粒的量呈正比的假设, 可以得到筛分动力学模型

$$E = 1 - e^{-kt} \quad (1.2)$$

式中  $E$ ——筛分效率;

$t$ ——筛分时间;

$k$ ——物料可筛性参数。

模型在假设条件下可以用一般数学方法推出, 但  $k$  与物料性质有关, 无法事先确定。对于特定的物料, 可以先进行一定数量的实验, 在实验数据的基础上, 再用数理统计的方法或单因素优化方法获得具体的  $k$  值。

### 1.2.2 根据模型中变量和时间的关系分类

(1) 稳态模型(静态模型)。稳态模型单纯反应生产过程变量之间的因果关系,不考虑时间的影响。一般用代数方程式来描述,是探求最优化的基础。

在实际生产过程中,当生产的条件波动幅度不大时,可认为是一个稳态过程,所观测的数据可看作是平均的数据,所建立的模型是稳态模型。目前选煤的一般计算都用稳态模型。

(2) 动态模型。生产过程中各变量的状态是随时间而变化的,此时各输入输出量之间的数学关系可以用微分方程或积分方程进行描述。

据分批浮选窄级别物料浮选动力学公式

$$d\gamma/dt = k(\gamma_{\infty} - \gamma)$$

积分整理得

$$\gamma = \gamma_{\infty}(1 - e^{-kt}) \quad (1.3)$$

式中  $\gamma$ ——精矿产率;

$t$ ——浮选时间;

$k$ ——浮选速率常数;

$\gamma_{\infty}$ ——最大回收率。

### 1.2.3 根据模型中变量的性质分类

(1) 确定性模型。自变量与因变量自身之间的关系都是确定的。

(2) 随机模型。全部或部分变量是随机变量,变量之间的关系不是确定性的函数关系,而是随机变化的相关关系。用数理统计方法建立的模型都是随机模型。随机模型建立后,往往还要检验它的可靠性和精确性。

根据模型的基本关系,数学模型还可以分成线性模型和非线性模型。根据变量的连续性,数学模型又可分成离散模型和连续模型。

虽然大多数问题是随机性的、动态的和非线性的,但由于确

定性、静态和线性模型容易处理，并且往往可以初步用来解决问题，所以建模时常先考虑确定性、静态和线性模型。

## 1.3 建立数学模型的方法和步骤

### 1.3.1 建立数学模型的方法和一般步骤

一般来说，建立数学模型的方法大体上可分成两大类：一类是机理分析方法，一类是测试分析方法。机理分析方法是根据对现实对象特性的认识，分析其因果关系，找出反应内部机理的规律。机理分析方法下建立的模型常有明确的物理或现实意义。该方法主要用于建立理论模型。测试分析方法将研究对象视为一个“黑箱”系统，其内部机理无法直接寻求，但可以测量系统的输入输出数据，并且以此为基础运用统计分析方法，按照事先确定的准则在某一类模型中选出一个与数据拟合得最好的模型，这种方法又称为系统辨识。该方法主要应用于经验模型。将以上两种方法结合起来也是常用的数学建模方法，用于建立综合模型。

建立数学模型的一般步骤，如图1-1所示。

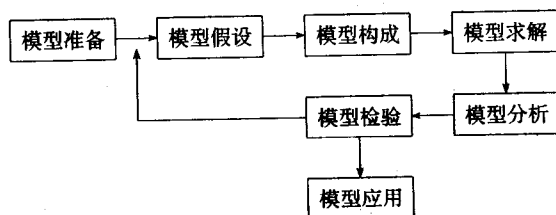


图1-1 建立数学模型的基本步骤示意图

(1) 模型准备。首先了解问题的实际背景，明确建立模型的目的，搜集建立模型必需的各种现象和数据等，尽量弄清对象的特征，由此确定用哪类模型。

(2) 模型假设。根据对象的特征和建立模型的目的，对问题进行必要的合理的简化。假设不合理或过于简单，会导致模型全

部或部分失败,此时应补充和修改假设;假设过于详细,试图将所有影响因素都考虑进去,可能使模型的建立变得复杂甚至无法继续下一步的工作。通常作假设,第一依据的是对问题内在规律的认识,第二依据的是对数据或现象的分析,也可以是两者的综合。作假设既要运用好相关的物理、化学、生物等方面的理论,又要充分发挥观察问题和分析问题的能力,要分清主次,抓住主要矛盾,尽量使问题线性化,均匀化。

(3) 模型构成。根据所作假设,分析对象的因果关系,利用对象的内在规律和适当的数学根据,构造各个量之间的等式、不等式关系或其它数学结构。

(4) 模型求解。利用解方程、绘图、证明定理、逻辑运算、数值计算等各种数学方法,特别是计算机技术,得到模型的具体形式。

(5) 模型分析。对模型解答进行数学上的分析,有时要根据问题的性质分析变量间的依赖关系或稳定状况,有时是根据所得结果给出数学上的预报,有时则可能要给出数学上的最优决策或控制。此外还常需要进行误差分析、模型对数据的稳定性或灵敏性分析等。

(6) 模型检验。把数学上分析得到的结果与实际的对象、数据进行比较,检验模型的合理性。模型检验的结果如果不符合实际情况,则问题通常出在模型的假设或模型构成上,应当修改、补充假设,重新建立模型。有些模型要经过反复比较,不断完善,才能获得较好的模型。

(7) 模型应用。根据问题的性质和建立模型的目的,将模型应用于实际对象。

### 1.3.2 建立经验模型的一般方法

在选煤数学模型中最常用的是经验模型,建立经验模型的具体步骤如下。

(1) 试验数据的整理。

获得数据是建立经验模型的先决条件。经验模型数据来源有两个途径,一个是生产上实际测量的数据,这类数据是生产过程的自然反映,获得这种数据的方法称为非可控试验法。另一个是将试验与建立模型的要求统一进行考虑,事先安排好试验来搜集数据,这种方法能用少量的试验得到建立模型所需的数据,并能简化建立模型过程中的数学计算,由于该方法的数据根据专门的试验来获得,所以称为可控试验法,或主动试验法。

因试验存在误差,数据的可靠性直接影响模型的准确性。建立模型时,首先要对数据进行检查和取舍。取舍标准来自专业知识或数理统计知识。如浮沉试验时如果出现高密度级灰分比低密度级灰分低时,则应重复试验。如果某一个试验重复多次,就可以利用数理统计的区间判断将误差较大的可疑数据舍去。本书的第2章将介绍数理统计的基础知识。

### (2) 模型形式的确定。

模型形式可以根据专业知识、实际经验和试验所取得的数据来决定。如果对某个过程的规律还不清楚,可以通过绘图勾画出数据的变化趋势,然后从熟悉的函数中选配。选煤过程中有四类曲线最常用:①粒度曲线和可选性曲线,一般从幂函数或指数函数中选配;②分配曲线,形状近似于S形曲线,可以选用S形函数;③速率方程,常见的浮选速率和筛分速率曲线,具有单调上升,逐渐趋于某一数值的特点,主要从指数函数中选配;④效率曲线,如浮选中精煤产率与药剂用量间经常存在一个极大值,这类曲线可以用抛物线函数表示。

### (3) 模型参数的估计。

模型形式确定后,其中的参数和常数还需确定,这个过程称为参数估计。参数估计可以采用回归分析或最优化等数学方法借助于计算机完成。本书第3章将介绍回归分析的主要内容,第5章将介绍最优化的基本方法。

### (4) 模型的检验。

模型建立起来后，能否真实地反映过程的特性，需要对模型进行检验。检验模型大部分是以模型的计算值和实测值相差多少为标准。在采用回归分析中还可以通过一些统计量来检验。模型显著并不意味着模型中所有的变量都显著。对于包含了多个变量的过程，还需对变量的显著性进行检验。

对实际问题，建立一个准确的数学模型是很不容易的，因而常常需要进行多次尝试，反复修正，逐步得到比较满意的模型。

## 2 数理统计基础

### 2.1 随机变量与分布

#### 2.1.1 随机过程与随机变量

在自然界中, 一个常见的现象是在一组相同的条件下, 一桩事件有时会发生, 有时不会发生, 这类事件就叫随机事件。相应的条件和结果统称为一个随机试验。一般设  $E$  为一试验, 以  $\omega$  表示它的一个结果, 称  $\omega$  为  $E$  的一基本事件, 所有基本事件的集合  $\Omega = \{\omega\}$  称为基本样本空间或样本空间。

如将硬币连抛 2 次, 观察硬币落下后是字面向上还是花面向上, 就是一个随机试验。以  $\omega_1$  代表字面朝上,  $\omega_2$  代表花面朝上, 于是有  $\Omega = \{\omega_1, \omega_2\}$ 。

设  $\Omega$  为某试验产生的样本空间,  $X$  为定义在  $\Omega$  上的实函数, 即对任一样本点,  $\omega \in \Omega$ ,  $X(\omega)$  为一实数, 则称  $X$  为一个随机变量。

如将硬币连抛  $n$  次, 记录硬币落下后是字面向上还是花面向上, 记  $X$  为字面向上的次数, 则  $X$  为随机变量。

对任意随机变量  $X$ , 称函数  $F(x) = P(X \leq x)$ ,  $(-\infty < x < \infty)$  为它的分布函数。任一随机变量的分布函数  $F(x)$ , 满足  $0 \leq F(x) \leq 1$ , 且具有以下基本性质: ① 单调不减。② 右连续。③  $F(-\infty) = 0, F(\infty) = 1$ 。

根据取值情况的不同, 随机变量可分为离散型随机变量和连续型随机变量。

(1) 离散型随机变量。如果随机变量  $X$  只能取有限个值, 并且以各种确定的概率取这些不同的值, 则称  $X$  为离散型随机变量。