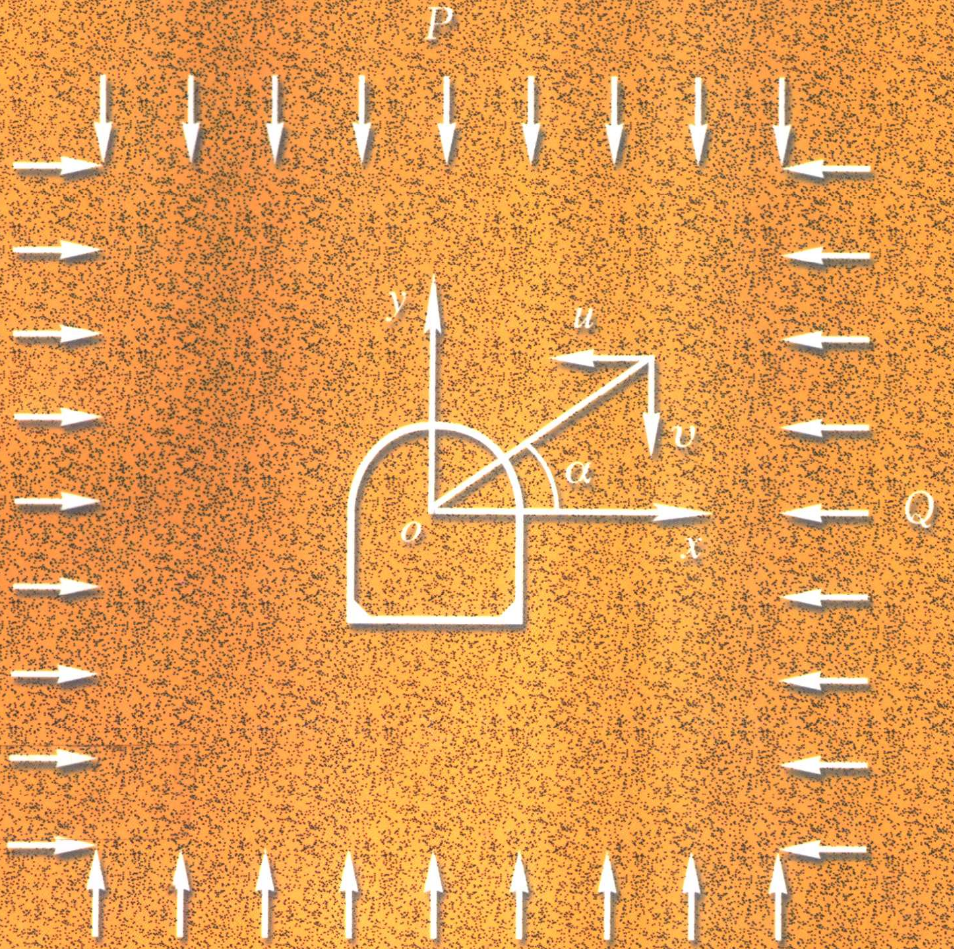
岩石力学反问题

吕爱钟　蒋斌松　著



煤炭工业出版社

994542

岩石力学反问题

吕爱钟　蒋斌松　著

煤炭工业出版社

内 容 提 要

本书系统阐述了岩石力学反问题的基本内容。全书共分十章：第一章介绍了反问题的内容及特点；第二章到第四章详细地介绍参数辨识的基础知识和简单线性模型参数辨识的各种方法，这是位移反分析的数学基础；第五章给出并证明参数可辨识条件，并对围岩弹性参数及地应力的可辨识性进行探讨；第六章介绍参数辨识方法在岩体位移反分析中的应用情况；第七章讨论金属支架荷载反算方法，对支架荷载反算解的稳定性和反算算法的合理选择等问题进行分析；第八章对位移反分析中用到的五种优化方法进行探讨，着重讨论这些优化方法在弹性、弹塑性和粘弹性位移反分析中的应用情况；第九章对位移反分析的测点最优布置和有限元网格划分范围等问题进行研究；第十章对岩石力学几何反问题——孔形优化进行论述，这部分内容为合理确定弹性岩体中地下巷道的最优开挖形状提供了理论基础；在附录中介绍矩阵、最优化方法、映射函数等有关数学基础知识。

本书可供铁道、矿山、水利水电及其他部门从事地下工程设计、施工和研究工作的技术人员参考，也可作为与地下工程有关专业的研究生的教材。

图书在版编目（CIP）数据

岩石力学反问题/吕爱钟，蒋斌松著. 一北京：煤炭工业出版社，1998. 7

ISBN 7-5020-1618-X

Ⅰ. 岩…　Ⅱ. ①吕…②蒋…　Ⅲ. 岩石力学-计算方法　Ⅳ. TU45

中国版本图书馆 CIP 数据核字（98）第 15567 号

岩 石 力 学 反 问 题

吕爱钟　蒋斌松　著

责任编辑：陈　昌　田克运

*

煤炭工业出版社　出版

（北京朝阳区霞光里 8 号　100016）

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 787×1092mm 1/16　印张 11 3/4

字数 272 千字　　印数 1—1，055

1998 年 8 月第 1 版　　1998 年 8 月第 1 次印刷

书号 4387　　定价 36.00 元

序

我怀着十分喜悦和赞佩的心情，粗读了山东矿业学院吕爱钟教授与蒋斌松副教授合著的《岩石力学反问题》一书，由衷感到该书的确是一本高层次、高水平的学术专著，应该尽早问世以飨读者。

1988 年中国岩石力学与工程学会在威海召开教育工作委员会期间，我曾阅读了吕爱钟同志赠送的《岩石力学反问题》讲义。那时在讲义中已经确立了岩石力学反问题的研究框架，认为位移反分析和孔形优化是岩石力学反问题的主要研究内容。当时他们已就位移反分析所涉及的一些重要问题开展了部分研究工作。

十年来，他们围绕岩石力学反问题中的很多重要问题又开展了进一步的深入研究，完成多项自然科学基金课题，取得一系列丰硕成果，发表多篇高水平的论文。该书就是在原讲义的基础上经过修改并补充上十年来的研究成果而完成的。

一份耕耘，一份收获，十年寒窗，换来了今天的硕果，实在可喜可贺。

岩石力学反问题的研究是促进岩石力学发展的重要途径，位移反分析可以用来确定原始地应力和岩体的力学性质参数，解决传统方法难以克服的困难；孔形优化可以用来选择地下硐室的最优开挖形状，以利于硐室的维护。

统观全书，与同类已出版的著作比较，具有以下特色和创新：

系统地讲述了参数辨识中的各种方法，为位移反分析奠定了理论基础；首次将参数可辨识条件应用于位移反分析；对位移反分析中的各种优化方法进行了全面深入的研究，这方面的工作在以往的位移反分析很少有人涉及；对位移反分析中的测点（线）最优布置和有限元网格划分范围等问题进行了系统研究，获得的成果为提高位移反分析精度提供了可靠的保证；在利用位移反分析确定地下结构荷载方面，对解的不稳定问题进行了全面的探讨，这在国内外尚属首次；地下硐室孔形的选择对围岩稳定性有相当大的影响，这是工程实践中已经知道的事实，作者在此方面开展了深入的理论研究工作，提出了新的孔形优化准则，研究成果可以用于确定地下洞室最优的开挖形状。

该书的出版为岩石力学反问题的深入研究奠定了坚实的基础，相信本书对岩石力学的发展将起到很好的推进作用。为此，我乐于写述了以上的一点文字，供请广大读者指正。

孙钧

1998 年仲夏于同济大学

孙钧：同济大学教授、中国科学院院士、国际岩石力学学会副主席、中国岩石力学与工程学会理事长暨国家小组主席。

前　言

本书所述岩石力学反问题，主要包括位移反分析和孔形优化两部分内容。

位移反分析在地下工程中主要有两方面的应用，一是通过现场量测的位移反求岩石力学参数和原始地应力，二是通过支架量测位移反算支架荷载。

本书第二章到第四章，讲述了参数辨识方法有关基础知识，这些内容是位移反分析的理论基础。

早期求解岩石力学问题的主要特点是把力学模型的选择、岩石性质参数及地应力的确定按三个过程单独进行。70年代末期位移反分析的出现，可以直接利用实地量测的位移值，根据选择的力学模型，同时求出岩石性质参数及原始地应力，本书在第六章中着重讨论了位移反分析在这方面的应用情况。

利用巷道围岩位移反求地应力及岩石性质参数是一种行之有效的方法，但这种方法并非能够求出所有参数，能求出的参数个数是一定的。作者在第五章深入地论述了这个问题，给出并证明了参数可辨识条件，首次将参数可辨识条件应用于弹性位移反分析中去。

位移反分析法从方法的实施过程来看，可分为正法和逆法两大类。逆法一般只适用于位移和待求参数为线性函数关系的情形，其特点可以将待求参数表示为量测位移的显函数，便于进行理论分析，此外还具有计算速度快等优点。正法比逆法具有更广泛的适用性，它既适用于位移是参数的线性函数的情形，也适用于位移是参数的非线性函数的情形。正法的另一个突出优点是可沿用现成的正问题计算方法及程序。最优化方法是求解正法的有力工具，目前已有多种优化方法用于位移反分析，以往的工作大都局限于用优化方法解决某一具体问题，很少在比较这些优化方法的优劣方面开展工作，所以对采用的优化方法是否可靠无法做出确切的回答。第八章通过弹性、轴对称弹塑性和粘弹性三种位移反分析多个算例的比较，从参数初始点的选择、收敛速度、收敛精度方面来评价这些优化方法的优劣。这方面的工作在以往的位移反分析中，很少有人涉及。

为了提高位移反分析的精度，在第七章十二节和第九章中详细地讨论了测点（线）的最优布置和有限元网格划分范围等问题，这些内容都是作者的研究成果。

在第七章给出了利用位移反分析确定地下结构荷载的方法——最小二乘和最小范数算法，这些方法都属于逆法。通过分析、检验发现：当反算荷载数目较多时，若采用常用的消去法求解逆法方程，会出现荷载反算解不稳定问题。作者详细分析了引起支架荷载反算解不稳定的原因，根据逆法方程的奇异程度，采用QR分解算法或奇异值分解算法求解逆法方程，以获得可靠的反算结果。

孔形优化是在力学模型及外荷载已知的情形下，寻找使孔洞周边或孔外域应力场满足预定要求的孔洞形状。它可以指导人们确定巷道（洞室）断面的开挖形状，使巷道在满足生产需要的工作空间要求下处于最佳的受力状态，以利于巷道的维护。地下巷道的稳定性与巷道的开挖形状密切相关，合理的开挖形状可以改善地下工程的受力状态，充分发挥围岩的自撑能力，以利于地下工程的维护，这是工程实践中已经知道的事实。第十章详细地

讨论了这类反问题，作者提出了一个新的孔形优化准则，可以保证用此准则所得孔形的孔边获得最小的应力集中。单从减少孔边应力集中的角度来看，这个准则是一个最佳的孔形优化准则，通过多个算例比较发现，它优于以往的其它孔形优化准则。通过算例还发现应力场的预定条件不能任意指定，若给出的预定条件不当，则不存在满足预定条件的孔形，也就是说此时反问题的解不存在，这有别于固体力学的正问题。

孔洞形状的优化离不开最优化方法，附录B给出了本书所利用的几种最优化方法。

本书的附录C和附录D也是作者近几年的研究成果。

作者于1987年为山东矿业学院硕士研究生编写了岩石力学反问题讲义。该讲义作为岩土工程专业研究生的教材已使用多届，本书就是在此初稿的基础上经过修改和十多年来研究成果的不断补充而完成的。

本书中的很多内容为国家自然科学基金和山东省、煤炭部自然科学基金项目的研究成果，这些项目都是在山东矿业学院陈子荫教授的主持下完成的。其中本书第七章部分内容借鉴和引用了陈子荫教授指导的硕士研究生（张中生、朱杰兵）的研究成果。

为了对岩石力学反问题有全面深入的了解，书中对国内外其他研究人员取得的成果也作了扼要介绍。

作者衷心感谢山东矿业学院陈子荫教授把作者带到岩石力学反问题的研究领域，并对冯豫教授对于完成本书所给予的关心和支持表示深深的谢意。

作者

1998年5月于山东矿业学院

目 录

第一章 绪 论

第一节 反 问 题

一、反问题的内容及特点

固体力学的正问题是指在物体的几何形状、材料性质及外荷载已知的情形下，求解物体内部的应力分布与变形规律；而相应的反问题是把问题中的某些待求量通过实地量测或人为指定变为已知量，而将某些已知量作为待求量。

近几年固体力学所关心的反问题有孔形优化问题，即已知材料性质及外荷载，设计孔洞的形状，使孔洞周边或孔外域的二次应力场（或位移场）满足预先提出的要求。

孔形优化问题也是岩石力学工作者所关心的一类反问题，它可以指导人们开挖巷道，使巷道在有一定工作空间的要求下处于最佳的受力状态，以利于巷道的维护。

矿山岩石力学所关心的第二类反问题是：已知巷道的开挖形状，根据实地量测的变形规律（位移场），求其描述这个系统的最佳模型及模型参数（岩石的性质参数）或原始地应力场，这类反问题称为位移反分析。

在数学中人们早就接触过反问题，例如，在初等代数学中，若把已知方程求根称为正问题的话，那么由根求方程的系数就是代数方程的反问题；在矩阵论中，由矩阵求特征值也对应着它的反问题——已知特征值反求矩阵。

由“结果”推断“原因”的反问题在人类认识自然与改造自然中起到了重要的作用，例如，遥测与遥感技术是通过接收回波（反射波）信息去判断人们感兴趣的物体的形状，地球物理勘探中的反问题就是借助于地球表面接收到的主动场或被动场的数据，经过处理判断地层的结构。

以往在矿山岩石力学中实际上也求解了一些反问题，例如，利用应力解除法求原始地应力，是通过量测应变或位移反求荷载，平板试验是利用量测位移反求岩体性质参数，所采用的求解方法都为逆法，只是未被普遍认识。

由“结果”推断的“原因”可能解不唯一（多解性），即某一特定“结果”可能引起的“原因”有多种，这是反问题的一类不适定性。反问题还可能具有解的不存在性和解的不稳定性这些特点。如果反问题的提法不正确，可能会导致反问题的解不存在。反问题的不稳定性是指实测资料有一定的微小误差时，反求出的结果产生很大偏差，甚至无法控制。如果反问题的解存在，唯一且稳定，则称反问题为适定的。

不适定问题的解法研究已成为计算数学的中心问题之一。在这一领域中，在理论上作出重要贡献的是原苏联学者古洪诺夫，1974 年他出版了“不适定问题的解法”一书，这是有关这方面的第一本专著，美国、中国相继翻译成英文、中文出版。

二、研究岩石力学反问题的意义

巷道形状优化设计（孔形优化）是一项很有实际意义的工作，它可以指导人们设计巷道断面，使巷道在有一定工作空间的要求下，处于容易维护的状态，达到既安全又经济的

目的。

孔形优化是在岩石性质参数及原始地应力已知的条件下进行的。岩石的性质参数及地应力的确定是解决岩石力学问题的关键所在，岩石力学工作者多年来一直在专门研究这个问题，但效果并不理想。岩石性质参数的确定一般都是在实验室或用现场试件进行的，试件尺寸与巷道尺寸比较仍然太小，试件不能反映实际岩体的结构，试件的受力状态与巷道的实际受力状态相差很大。这样，根据试件确定的岩石性质参数解决实际的岩石力学问题，其误差很大。

计算结果与实际量测结果相差很大并不是完全由以上原因引起的。通常，原始地应力测定可靠性较差或者是选择的力学模型不正确都可以造成很大的误差。

以往求解岩石力学问题的主要特点是把力学模型的选择、岩石性质参数及地应力的确定三个过程单独进行。现在，采用位移反分析，我们可以直接利用实地量测的变形规律，根据选择的力学模型，同时求出岩石性质参数及原始地应力。

实地量测是一种最反映实际情况的现场试验，究竟采用何种力学模型这不能凭空事先决定，而必须由实地量测的变形规律在已知的一组模型里求出与实际变形规律最接近的最佳模型，这就是模型鉴别的内容。求出了模型（包括参数）和原始地应力，我们再按正问题去计算，预测以后开挖所表现的各种力学行为。

确定支护结构上的荷载，这是地下结构设计与地面结构设计的最大不同点。地面结构所承受的荷载较易确定，而地下支护结构所承受的荷载是不能事先知道的。地下支护结构承受的荷载取决于结构与岩体的相互作用，它的大小及分布规律与岩体性质、原始地应力场、支护刚度及支护时间等多种因素有关。分析结构与岩体的相互作用，要利用岩体和结构的力学模型，但由于目前岩体力学模型的研究尚未成熟，故根据相互作用从理论上精确给出结构上的荷载是困难的。现在有了位移反分析方法，我们可以在可靠性较高的结构模型的基础上，利用结构上的位移量测值反求结构上的荷载。当结构的力学模型正确时，反算的荷载是可信的。

第二节　系统辨识和参数辨识

如果把所讨论的对象作为一个系统的话，则正问题是指已知描述系统的模型及输入，求输出，如图 1. 1 所示。在这种情况下，不但模型结构是已知的，而且所有有关的参数也是已知的。

图 1.1　正问题的求解过程

而反问题是指通过量测输出，求系统的模型或模型参数。在有些情况下，当模型和模型参数已知时，反问题是指由输出求输入。

按对系统的了解程度，反问题可分为系统辨识和参数辨识两类。

一、系统辨识

系统辨识是通过量测得到系统的输出和输入数据来确定描述这个系统的数学方程，即模型结构。为了得到这个模型，我们可以用各种输入来试探该系统并观测其响应（输出），

然后对输入——输出数据进行处理来得到模型。

近年来，系统辨识的应用领域日益扩大，在通信工程、航空航天工程、地质学、经济学、生物学、医学等方面都得到应用，各个领域都在利用系统辨识方法建立各自系统的定量模型，从而由定性到定量地解决实际问题。另一方面，也由于现代计算工具发展，使许多问题可以通过计算机加以解决，这又推动了系统辨识的发展。

基于对系统先验信息的了解程度，我们可以把系统辨识问题分为两类：

“黑箱问题”，也叫完全辨识问题。在这种情况下，被辨识的系统的基本特性是完全未知的。例如，系统是线性的还是非线性的，是动态的还是静态的，对这些基本的信息都一无所知，要辨识这类系统当然是很困难的，目前尚无有效的办法。

“灰箱问题”，又叫不完全辨识问题。在这一类问题中，系统的某些基本特性（例如线性）为已知，不能确切知道的只是系统方程的阶次和系数。当然，这类问题比“黑箱问题”容易处理。

幸好，许多工程上的辨识问题属于“灰箱问题”，这样，系统辨识问题就简化为模型鉴别和参数辨识问题了，参数辨识是系统辨识中最重要也是研究得最成熟的部分。

二、参数辨识

参数辨识是在模型结构已知的情形下，根据能够量测出来的输入和输出，来决定模型中的某些或全部参数，如图 1.2 所示。

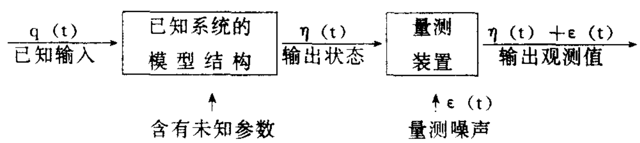


图 1.2　参数辨识过程

由于量测总含有误差，故求出的参数值只是真实参数的估计值，利用概率统计的知识来辨识参数就是参数估计问题，所以在很多文献中把参数辨识称之为参数估计（Parameter-estimation）。

参数辨识是近几年发展较快的年轻学科，在各个领域都引起了重视，它的名字还没有完全统一起来，参数辨识的其它名字有非线性估计（nonlinear estimation）、非线性回归（nonlinear regression）、参数优化（optimization of parameters），有的文献干脆称为建模（model building）或系统辨识（identification of systems）。

“估计”是数理统计中的术语，“辨识”是电气工程上的术语。

对于矿山岩石力学问题，一般把易量测的位移作为系统的输出，巷道及支护的形状、尺寸作为输入，与模型结构有关的变形参数可作为模型参数，地应力既可以作为输入，也可看作为待识别的参数。

第二章　参数辨识方法的基础知识

第一节　参数辨识的几个要素

一、模型

在自然科学和工程领域中，模型的建立与实验、观察具有同等重要的地位，模型的建立是实验、观察、认识问题的一个飞跃。模型是实际系统“原型”的一种“类似”，它与“原型”必定存在一定的差别。任何原型都有数不清的层次和特征，能反映出原型一切特征的只能是原型本身，而不是模型。建模的目的不是将原型的一切方面都表达出来，模型只需在所要研究的主题范围内能表达人们最需要知道的那些特征即可，从而达到对原型的抽象。以模型为基础，较方便地对原型进行分析、研究，以便通过模型的预测结果来正确指导我们作出某种决策。

模型的表达形式可以是概念性的、物理的或者是数学的，这取决于模型建立的特定目的。采用数学描述的形式所建立的模型称为数学模型，它是系统中的各个物理量之间的关系所构成的数学结构，象代数方程、微分方程等等。不言而喻，目前在岩石力学中采用的弹性模型、弹塑性模型、粘弹模型都为数学模型。

模型结构的形式有：静态的或是动态的，线性的或是非线性的，参数是定常的或是时变的，确定型的或是随机型的，参数模型或是非参数模型。

二、参数和状态

由常微分或偏微分方程给出的数学模型，有时它的解是一组比较简单的代数方程。在任何情形下，都有自变量和因变量以及一些常数。因变量有时称为状态变量（或信号），而常数称为参数。

在实验中，常常直接量测的是状态，而参数一般不能直接量测出来，参数只能由状态的量测值反求出来。有的教科书所关心的是参数估计问题，而有的教科书则侧重于状态的估计（预测）。参数估计与状态估计（预测）两个问题非常相似，在参数估计的同时，通常状态估计（预测）自动完成。

参数和状态这两个概念可以用下面的简单例子说明。

例 1：

根据牛顿第二定律可知：　$a_x(t) = \frac{1}{m}F_x(t)$

这里 $F_x(t)$ 是 x 方向的力，$a_x(t)$ 是 x 方向的加速度，它们都是时间 t 的函数，m 是质量。可把 $F_x(t)$，$a_x(t)$ 看作为状态，质量 m 则是参数。力和加速度通常可容易地通过量测获得。对此问题，质量不但可以根据力和加速度求得，而且也可以直接量测获得，但是对有些情况，质量则必须根据力和加速度推算而得，例如，确定慧星和行星的质量，人们不可能直接量测它们的质量。

例 2：

以初速度v_0垂直上抛一个物体，已知物体离开地面的距离s可由公式$s=v_0t-\frac{1}{2}gt^2$表示。这里g是重力加速度，它是一个参数，时间t为自变量，s是状态。v_0既可看作为参数，也可看作为状态。

例 3：

一等截面拉杆，截面积为A，原长为L，它一端固定，一端受拉力P的作用(图 2.1)，每个截面都产生x方向的位移，距原点O为x处的截面位移为$u(x)$：

$$u(x)=\frac{Px}{EA}$$

这里u (x) 为状态，x是自变量，E、A、P为参数，但A、P可直接量取获得，E必须由状态值求得。

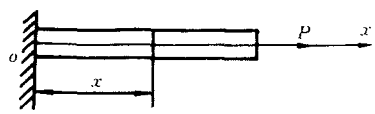


图 2.1　等截面杆受拉力P的作用

以上所举的这些参数与统计参数比较，一般称为物理参数，量测误差的方差、相关系数这样的参数称为统计参数。物理参数和统计参数在某些问题中可能都要辨识，而我们最关心的是物理参数的辨识。

三、准则函数

若模型能精确地反映我们对“原型”所关心的那些特征，则模型的输出就是系统的实际输出。如果对输出的量测值也不存在误差，且所讨论的反问题为适定的，则由量测的输出总可列出也只能列出与待辨识参数个数相等的独立方程，由这些方程即可唯一地求出待求的参数。

实际上，由于模型的近似性和量测误差的存在，则按以上方法求得的参数不能很好地反映整个系统的特征。如何能够求出反映整个系统的最优参数呢？最直观的做法是：量测的数量必须大大地超过待求参数的个数，这样可以降低量测噪声对待求参数的影响，这样列出的方程个数多于待求的参数个数，所得的方程组为矛盾方程组，通过适当的最优化技术可以求解这样的问题，使得在某种意义下求得的参数为最优参数。如何衡量最优？最优准则如何确定？这是参数辨识首先要解决的问题。

一般把最优化准则称为准则函数，记为J。准则函数总体上可分为两大类，一类是以输出信号为基础的准则函数，一类是以量测误差或参数的概率统计性质为基础的准则函数，后面我们将分别称它们为第一类和第二类准则函数。

第一类准则函数，一般表示为系统的实际输出量测值$y(t)$和模型的输出$\eta(t)$的偏差的某个函数，例如，可取

$$J=\sum_{i=1}^{n}[y(t_i)-\eta(t_i)]^2 \tag{2-1}$$

等一些误差函数作为准则函数。式中n为量测数量，$\eta(t_i)$是输入和参数的函数，给定模型结构也就是知道了$\eta(t_i)$的函数形式，t是自变量。对于以时间作为自变量的模型，t_i表示第i时刻，对于以位置作为自变量的模型，t_i表示第i个位置。

$y(t_i)$是已知的量测值，当输入为已知时，显然，准则函数J的大小随着所选的模型参数不同而不同，当J达到最小值时的参数即为最优参数。

对于这类准则函数，参数辨识实际上可作为一个最优化问题处理，即通过所选的准则函数寻求使准则函数达到极小的参数值。就此而言，准则函数称为目标函数。根据求解的

问题不同，在不同场合下J往往还有其它的名字，例如误差函数、损失函数、成本函数等等。

以量测误差或参数的概率统计性质为基础的第二类准则函数的参数辨识，事先考虑了输出信号量测误差的统计特性，把待求参数作为确定性常数或随机变量。参数的最优并不是象第一类准则函数直接以输出的偏差最小为衡量准则，而是以参数误差（参数真值与参数估计值的差）为最小或以特定输出量测值出现可能性为最大等概率统计特性为衡量准则。

对于第二类准则函数，参数辨识作为估计问题处理，参数估计的具体实现同样离不开最优化技术。

两类准则函数相比，由于后者利用了一些概率统计知识，所以后者比前者的最大优点是可以计算量测噪声对参数辨识的影响程度，有时所求出的参数估计值具有较好的统计特性。

第二节　参数辨识的方法分类

参数辨识具有多种方法。

根据不同的准则函数可得出一系列参数辨识法，例如，以第一类准则函数为基础的最小二乘法、加权最小二乘法，以第二类准则函数为基础的最小方差法、极大似然法、贝叶斯法等等。以第一类准则函数为基础的各种参数辨识方法统称为确定性方法，以第二类准则函数为基础的各种参数辨识方法统称为随机性方法。

根据辨识的方式可分为离线辨识和在线辨识，所谓离线辨识是在全部量测数据的基础上求解模型参数；而在线辨识是指收集到新的量测值以后，就在前一次参数估计值的基础上立刻进行递推计算，尽快地给出新的估计值，例如，序贯最小二乘法就是一种在线辨识。

准则函数选定以后，参数辨识的过程是寻求准则函数的极值点。对于岩土工程问题，根据问题的性质及寻求准则函数极值点的算法，参数辨识方法可分为逆法和正法两大类。

现以第一类准则函数式（2—1）为例说明逆法和正法。

所谓逆法是指能把模型输出表示成待求参数的显函数，由模型输出的量测值，利用这个函数关系反求出待求参数，这个过程恰好与正问题的求解过程相反，逆法由此而得名。

若考虑误差的存在，则这个方法的具体实施过程是这样的：由式（2－1）的 J 达到极小值的必要条件（即 J 关于参数的一阶偏导数等于零）所列的方程组，求解待求参数。

逆法的应用决定于模型输出是否能表示成待求参数的函数，以及函数关系的性质如何？

当模型输出是参数的线性函数时，一般才利用逆法。此时，由极值必要条件所列的以参数为未知量的方程组为线性方程组，线性方程组的求解是轻而易举的事。

模型输出可以是待解的微分方程组，也可以是它的解析解。若已有解析解，参数辨识就可以以解析解为出发点，这时计算量少。若无解析解则需要以待解方程组及具体问题的定解条件为出发点，在岩土工程中，这是一种普遍情形，这时要利用有限单元法等数值方法。

正法与逆法不同，它不是利用极值的必要条件求出待求参数，而是首先对待求参数指定“初值”，然后计算模型输出值并和输出量测值比较。如果吻合良好，假设的参数“初值”就是要找的参数值。实际上当然不会这么巧，这时修改参数值，重新计算模型输出值，重新比较一直到准则函数达到极小值，此时的参数值即为所要求的参数。

若模型输出是待解的微分方程组，则由参数“初值”计算模型输出是求解正问题，由

此看出此时的参数辨识过程是解一系列正问题，正法由此而得名。

可以看出，正法和逆法都是寻求准则函数的极小点，但寻求的算法不同。正法比逆法具有更广泛的适用性，它既适用于模型输出是参数的线性函数的情形，也适用于模型输出是参数的非线性函数的情形。它的另一个优点是仍可沿用现成的正问题计算方法及程序。

最优化技术中的直接法是求解正法的有力工具，模式搜索法（Hooke－Jeeves）（也称步长加速法）、变量轮换法、单纯形法、鲍威尔（Powell）法等方法都是最优化技术中广泛应用的直接法。

以上用第一类准则函数说明了正法和逆法，同样对第二类准则函数也存在正法和逆法。

第三节　参数辨识方法

本节主要讲述按第一、二类准则函数分类的各种参数辨识方法的基本思想，而不涉及方法的本身细节。

一、观测变量和参数之间的关系

一组可观测变量（$y,x_1,\cdots,x_k$），一批参数（$\beta_1,\cdots,\beta_m$）和一组随机变量（$\varepsilon,\varepsilon^{(1)},\cdots,\varepsilon^{(p)}$）总可以假定以一定的函数关系存在。这里，我们只讨论每次观测只包含一个随机变量的情形，即只有一个 ε 存在，其它的 $\varepsilon^{(1)},\cdots,\varepsilon^{(p)}$ 不考虑，不同次的观测用下标 i 区别，即第 i 次的观测包含的随机变量用 ε_i 表示。

选择一个量测变量 y，并用其它量测变量 $x_1,\cdots,x_k$ 表示，即

$$y = f(x_1,\cdots,x_k;\beta_1,\cdots,\beta_m;\varepsilon) \qquad (2-2)$$

通常称 y 为因变量，称 $x_1,\cdots,x_k$ 为自变量。

如果误差 ε 是可加的，即

$$y = \eta(x_1,\cdots,x_k;\beta_1,\cdots,\beta_m) + \varepsilon \qquad (2-3)$$

则对于求解问题是方便的。在式（2－3）中，ε 的分布与未知参数 $\beta_1,\cdots,\beta_m$ 无关。

把式（2－3）用矢量记法进行缩写对于书写是方便的，记

$$\boldsymbol{X} = (x_1,\cdots,x_k)^{\mathrm{T}}$$

$$\boldsymbol{\beta} = (\beta_1,\cdots,\beta_m)^{\mathrm{T}}$$

则

$$y = \eta(\boldsymbol{X};\boldsymbol{\beta}) + \varepsilon$$

第 i 次观测通过下标 i 表示，即

$$y_i = \eta(x_{i1},\cdots,x_{ik};\beta_1,\cdots,\beta_m) + \varepsilon_i = \eta(\boldsymbol{X}_i;\boldsymbol{\beta}) + \varepsilon_i = \eta_i + \varepsilon_i$$

进行了 n 次量测，则有 $y_1,\cdots,y_n,\varepsilon_1,\cdots,\varepsilon_n$ 等等，为了书写更方便，则可记

$$\boldsymbol{Y} = (y_1,\cdots,y_n)^{\mathrm{T}} \quad \boldsymbol{\varepsilon} = (\varepsilon_1,\cdots,\varepsilon_n)^{\mathrm{T}}$$

（$\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n$）可写成矩阵形式，即

$$\boldsymbol{X} = [\boldsymbol{X}_1,\cdots,\boldsymbol{X}_n]^{\mathrm{T}} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}$$

ε 的分布一般来说是未知的，如果 $\varepsilon_i(i=1,\cdots,n)$ 是相关的，那么以第二类准则函数为基础的参数辨识困难较大，本文主要讨论 $\varepsilon_i(i=1,\cdots,n)$ 不相关的情形。

二、线性问题

如果能够把 $\eta(\boldsymbol{X};\boldsymbol{\beta})$ 写成以下形式

$$\eta(\boldsymbol{X};\boldsymbol{\beta}) = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m \qquad (2-4)$$

即在同一时刻（或同一位置），模型输出 $\eta(\boldsymbol{X};\boldsymbol{\beta})$ 是参数 $\boldsymbol{\beta}$ 的线性函数，则称式(2－4)为线性模型。

对于线性模型，有：

$$y = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_m x_m + \varepsilon \qquad (2-5)$$

由式(2－5)来确定参数 $\boldsymbol{\beta}$ 的估计值，称为线性估计。

如果模型输出 $\eta(\boldsymbol{X};\boldsymbol{\beta})$ 是参数 $\boldsymbol{\beta}$ 的非线性函数，则称非线性模型，非线性模型的参数估计比线性模型的参数估计复杂。

三、常规最小二乘估计

已有一组量测值 $y_1,\cdots,y_n$，则在模型结构已知的情形下可写出：

$$J = \sum_{i=1}^{n}[y_i - \eta(\boldsymbol{X}_i;\boldsymbol{\beta})]^2 \qquad (2-6)$$

使(2－6)达到极小值的 $\hat{\boldsymbol{\beta}}$ 称为参数估计值。这个参数估计方法称为常规最小二乘法（ordinary least squares method）。对于线性模型，由逆法可以求出 $\hat{\boldsymbol{\beta}}$ 的解析式，对于非线性模型，一般用正法求解 $\hat{\boldsymbol{\beta}}$。

常规最小二乘法没有利用 ε 的概率统计特性。

四、高斯—马尔可夫估计（Gauss－markov estimation）

如果对于不同的量测时刻(或量测位置)，式(2－5)中的 $\varepsilon_i(i=1,\cdots,n)$ 是不相关的，但 ε_i 的方差不尽相同，若令 ε_i 相应的方差分别为：$\sigma_1^2,\cdots,\sigma_n^2$，则可令准则函数为：

$$J = \sum_{i=1}^{n}\frac{(y_i - \eta_i)^2}{\sigma_i^2} \qquad (2-7)$$

由式(2－7)极小求出的参数估计值 $\hat{\boldsymbol{\beta}}$ 一般比常规最小二乘法求得的参数估计值好，这是因为量测精度越高，相应的量测数据占的比重越大，这由式(2－7)可清楚看出(量测精度越高，误差 ε_i 的方差 σ_i^2 越小)。

五、最大似然估计（maximum likelihood estimation）

如果已知 $\varepsilon_i(i=1,\cdots,n)$ 的联合分布形式，那么 $y_i(i=1,\cdots,n)$ 的联合分布即可确定，此时把参数作为未知的确定性常量，则 y_i 的联合分布为条件分布，它可由概率密度函数 $p(\boldsymbol{Y}/\boldsymbol{\beta})$ 描述。

由 $p(\boldsymbol{Y}/\boldsymbol{\beta})$ 极大求出的参数估计值 $\hat{\boldsymbol{\beta}}_{ML}$ 称为参数 $\boldsymbol{\beta}$ 的最大似然估计，$p(\boldsymbol{Y}/\boldsymbol{\beta})$ 即为准则函数。

最大似然估计与 $\varepsilon_i(i=1,\cdots,n)$ 的分布有关，后面将会发现，如果 ε_i 是零均值同方差的独立正态分布(高斯分布)时，最大似然估计量与常规最小二乘估计量是相同的。

六、最大验后估计与贝叶斯估计

如果随机变量 ε 的分布已知(并知道 ε 的分布参数，例如方差是已知的)，待求参数 $\boldsymbol{\beta}$ 也为随机变量，并且它的概率密度函数 $p(\boldsymbol{\beta})$(称为参数 $\boldsymbol{\beta}$ 的验前分布）也为已知，那么根据贝叶斯定理，我们可以获得最大验后估计（MAP estimation，MAP 是 Maximum A Posterior 的简写）和贝叶斯估计（Bayes's estimation）。

最大验后估计量 $\hat{\beta}_{MAP}$ 是根据验后分布 $p(\beta/Y)$ 取极大值的条件求得的，它是观测值为 Y 的条件下，参数 β 的“最可能”数值，$p(\beta/Y)$ 即为准则函数。

为了方便地说明贝叶斯估计，下面讨论待估参数只有一个分量的情形。

设 $\hat{\beta}$ 为 β 的估计值，则定义准则函数为：

$$J = \int_{-\infty}^{+\infty} (\hat{\beta} - \beta)^2 p(\beta/Y) d\beta \tag{2-8}$$

使上式达到极小的 $\hat{\beta}_B$ 称为贝叶斯估计。由式(2 － 8) 可得贝叶斯估计量 $\hat{\beta}_B = \int_{-\infty}^{+\infty} \beta p(\beta/Y) d\beta$，这就是参数验后分布 $p(\beta/Y)$ 的期望值。

式(2－8)表示参数估计值 $\hat{\beta}$ 对真值 β 的均方差，所以在取得观测值 Y 的条件下，平均来讲，随机参数 β 最靠近其贝叶斯估计值 $\hat{\beta}_B$。因此，可以说 $\hat{\beta}_B$ 是参数 β 的“最有效”的估计值。

最大似然估计、最大验后估计、贝叶斯估计与最小二乘估计、高斯 — 马尔可夫估计一样，既适合线性模型的情形，也适合非线性模型的情形。但对于非线性模型，难于直接求出概率密度函数 $p(Y/\beta)$ 或 $p(\beta/Y)$，所以此时的参数估计也存在一定的困难，它们可通过线性模型的估计值经多次迭代而求得。

当模型结构和参数 β 的验前分布满足一定的要求时，最大似然估计量、最大验后估计量、贝叶斯估计量与最小二乘估计量、高斯 — 马尔可夫估计量是相同的。

第四节 参数估计量的统计特性

上节介绍的各种参数辨识方法，一种情形是事先把参数作为确定性的量，一种情形是事先把参数作为随机变量。两种情形求出的参数估计量都是随机变量，后种情形是显然的，前种情形也不难理解。由 n 次量测可求出参数估计值 $\hat{\beta}$，由于量测误差的存在，所求出的估计值不可能是参数真值，如果另取一组 n 个量测值去求参数，由于误差的随机性，则得到的 $\hat{\beta}$ 和上次不同，所以说，尽管参数 β 本身不是随机变量，但是这样求出的估计量 $\hat{\beta}$ 却总是随机变量。

估计量是随机变量，必须从统计的观点分析，衡量估计量的优劣，无偏性、有效性和一致性是鉴别和比较估计量好坏的重要标准。

一、无偏性

人们总是希望未知参数 β 与它的估计值 $\hat{\beta}$ 在某种意义上最接近，当然作为某一次的估计值它与真值可能是不同的，然而，如果通过一系列试验求出的不同估计值，我们很自然地要求这些估计值的平均值与未知参数的真值相等，这就是说，要求参数 β 的估计量 $\hat{\beta}$ 的数学期望(也有称数学期望为均值)等于参数的真值 β，即如果关系式

$$E(\hat{\beta}) = \beta$$

成立，那么称满足这种要求的估计量 $\hat{\beta}$ 为参数 β 的无偏估计量。

估计量的无偏性意味着无论重复多少次量测，要求估计值都能在被估计参数的真值附近摆动，而其平均值就等于参数真值。

二、有效性

无偏性还不能完全决定估计量的性质，对于一个估计量，还需要进一步考虑到估计值和参数真值的平均偏离的大小，或者说估计值围绕真值摆动幅度的大小问题。方差能够反