

成都理工大学

博士学位论文

地球物理反演中病态矩阵方程正则化解算方法研究

姓名：王文娟

申请学位级别：博士

专业：地球探测与信息技术

指导教师：曹俊兴;谭永基

20100501

地球物理反演中病态矩阵方程正则化解算方法研究

作者简介：王文娟，女，1970 年 05 月生，师从成都理工大学曹俊兴教授，复旦大学谭永基教授，2010 年 06 月毕业于成都理工大学地球探测与信息技术专业，获得工学博士学位。

摘 要

地球物理反演是地球物理探测数据最重要的解释方法技术。求解地球物理反演问题常会涉及到大型病态矩阵方程的求解。理论上讲正则化方法是处理病态问题的有效手段，但在实践上正则化参数的选择却是一个困难的问题。本文在比较系统地研究了正则化方法的基础上，针对实际计算中常会碰到的问题，将其与 Active-Set 算法、差分进化算法等相结合，发展了一些新的病态矩阵方程正则化解算方法。论文的主要内容包括：

1. 研究了在实际反演中遇到参数有非负要求特性时的反演计算方法。将原问题转化为一个带非负约束的阻尼最小二乘问题，并用 Active-Set 算法求解。通过对理论模型进行数值模拟计算，验证了将 Tikhonov 正则化方法与 Active-Set 算法相结合的 A-TR 算法的有效性。应用到实际双频电导率成像反演，也取得了满意的结果。

2. 研究了差分进化算法在地球物理反演中的几种应用。为加速差分进化算法的收敛速度，提出了将种群熵的自适应差分进化(ARDE)算法以及粒子群差分进化混合(PSODE)算法分别与 Tikhonov 正则化方法结合。在大型反演计算中，这两种方法可以在不影响反演效果的前提下，不同程度地提高收敛速度，降低时间成本。同时结合 LSQR 和差分进化算法的优点，提出了基于 LSQR 算法的差分混合(HDE)算法，避开了 Tikhonov 和 TSVD 等直接正则化算法在选取正则化参数上的困难，同时具有数值稳定性好、不依赖于初值、不易陷入局部极值和收敛速度快等优点，适宜于在正则化参数选取困难情况时的地球物理反演问题的求解。

3. 提出了一种双参数混合正则化方法。引入了带有二阶正则算子的正则化项，并应用 L-曲线法、偏差原理和广义交叉校验准则的优化组合确定了最佳正

则化参数. 数值模拟实验和实际数据处理实验结果表明了该方法的可行性. 这是一种将高阶正则化算子应用于实际反演计算的新的尝试.

基于数值模拟实验和实际数据处理实验, 认为研究发展的 **A-TR** 算法、**HDE** 等算法各有其不同的适用条件, **A-TR** 算法适用于求解反演参数有非负约束的情况, 而当正则化参数选取困难时, 可采用 **HDE** 算法. 针对本文所考察的双频电导率反演问题, 由于电导率的非负性, 采用 **A-TR** 方法可得到更加精细可靠的重建图像.

关键词: 地球物理反演 正则化方法 差分进化算法 正则化参数

Regularization Algorithms for Solving Ill-posed Matrix Equations Arising from Geophysical Inversions

Introduction of the author: Wangwenjuan, female, was born in May, 1970 whose tutor was Professor Cao Junxing and Professor Tan Yongji. She graduated from Chengdu University of Technology in Earth Exploration & Information Techniques major and was granted the Doctor Degree in June, 2010.

Abstract

Geophysical inversion is a key technique in geophysical exploration. Geophysical inversion often relates to solving ill-posed, large matrix equations. Theoretically, regularization is an effective method in dealing with ill-posed problems. Its application in practical geophysical inversion problems, however, still has many difficulties in selecting regularization parameters. Based on systematic study of regularization, this dissertation presents a few newly developed regularization methods that apply Active-Set algorithm, Differential Evolution algorithm, and a few others. These regularization methods focus on solving ill-posed matrix equations arising from practical problems. Main results of the dissertation include:

Tikhonov regularization and Active-Set algorithm are applied together to the geophysical inversion problems so that the problems with non-negative parameters are converted into problems of non-negative damped least square algorithm, which can be further solved by the Active-Set algorithm. The improved recursive algorithm is further verified by numerical simulation. Satisfactory results are obtained by applying this algorithm to electrical conductivity imagery inversion.

Furthermore, differential evolution algorithms are also studied. To improve the rate of convergence of Differential Evolution algorithms, two new Tikhonov regularization algorithms are proposed that respectively employ Adaptive Recursive Differential Evolution (ARDE) algorithm based on population entropy and Particle Swarm Optimization and Differential Evolution (PSODE) algorithm. Without any compromise in effectiveness, these two algorithms both improve the convergence speed and thus reduce computation cost. Still further, a new DE algorithm based on LSQS, which inherits the advantages of both LSQR and DE, is designed. This new algorithm avoids the common difficulty of regularization parameter selection in Tikhonov and TSVD algorithms. It also displays superior stability, independence on initial values, unlikelihood of local extrema, and faster convergence. This algorithm is particularly suitable to solve the problem of selection of regularization parameters in the study of geophysical inversion.

Finally, the selection method of regularization term is also studied. A regularization term with a second order regularization operator is introduced to

propose a mixed regularization method with two parameters. The L-curve criterion, discrepancy principle, and generalized cross-validation are applied to determine the optimal value of the regularization parameter. The validity and superiority of the proposed method is verified by numerical simulation of the theoretical model.

Based numeric simulations and results from actual data processing, it is found that the newly developed A-TR regularization algorithm, HDE regularization algorithm are effective in certain conditions. The A-TR regularization algorithm is applicable to inversion problems that require non-negative parameters, whereas HDE regularization algorithm is applicable to those inversion problems whose selection of regularization parameters is difficult. Because of non-negative conductivity, A-TR algorithm will get more detailed and reliable imagery for dual-frequency conductivity inversion problems investigated by this paper.

Keywords: Geophysical inversion; Regularization method; Differential evolution algorithm; Regularization parameter.

独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。据我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得成都理工大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的人员对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：

年 月 日

学位论文版权使用授权书

本学位论文作者完全了解成都理工大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权成都理工大学可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后适用本授权书）

学位论文作者签名：

学位论文作者导师签名：

年 月 日

第1章 引言

1.1 选题依据

自上世纪六十年代以来,“由效果、表现(输出)反求原因、原像(输入)”^[1]的“数学物理反问题”在众多的科学工程技术领域,如地球物理、遥感技术、模式识别、大气科学以及金融科学、经济决策等管理工程领域中出现.这类问题有着广泛的应用背景,为此在国际上引起越来越多科学家的兴趣和重视.

数学物理反演问题来源于社会发展和科学技术的驱动.1902年 Cauchy 第一次提出反演问题的时候,人们对此曾不屑一顾.但随着科技的发展,人们发现越来越多的问题都涉及求解反演问题,如地球物理观测数据的反演、遥感参数反演等等.

从实际应用角度来看,概括地说,有两个不同的动机驱动着反问题的研究^[1]:一方面是希望了解物理过程过去的状态,辨识一些有用的物理参数;另一方面是希望通过对当前状态的了解,能够去调整、干预某些参数,从而达到影响、控制一个系统的作用.

地球物理学家一直致力于求解反演问题.但是地球物理成像反演所涉及的大型线性方程组通常呈现不相容性,只能在某种意义上求其近似解.这就是地球物理中常见的病态方程在数学上的一个共同的特点.由于病态方程的病态性,它们的解不连续依赖于观察数据.为了解这类病态方程,许多学者进行了深入的研究,提出和发展了各种有效的数值方法.如奇异值法(简称 SVD)、截断奇异值法(简称 TSVD)、最小二乘法、数值相关性技术^[2]、条件预优和预条件子方法^[3,4]等,但最完善和有效的还是前苏联科学院院士 Tikhonov 提出的正则化方法^[5]以及在此基础上提出的一些改进的正则化方法.

近年来由于最优化方法的发展,人们越来越多地关注把此类方法结合并应用到反演问题的计算求解上,取得了一系列应用于实际的优秀科研成果.对地质勘探,应用这些方法可以较准确地确定不为人知的地下情况;对图像恢复问题,应用这些方法可以在很大程度上抑制噪音和误差的传播,从而达到稳定的恢复效果;对金融科学中的市场波动率问题,应用这些方法可以较科学地估算出一定时间段内的股票市场波动率,从而可以为投资者提供较准确的决策支持;对遥感地表参数反演和大气参数反演问题,利用正则化方法和最优化方法,可以精确地获得地表反照率和预测在一定时间和空间范围内大气(如气溶胶)的变化状况,从而为气象局、环保局、海洋局等国家决策部门提供决策依据.由此可以看出,利用正则化方法和最优化方法的有效结合所提出的新型算法,在地质地貌勘探等众多的科学技术领域具有广泛的应用价值^[1].

1.2 国内外研究现状

最近三十年来,对数学物理反演问题的研究进展十分迅速,已经成为应用数学中的一个热点前沿研究领域.该学科的发展,在很大程度上受其他学科与众多工程领域的迫切需求所驱动.反演问题具有涉及面广、内容丰富、跨行业、跨学科等特点.从反演问题的研究方法上看,它用到应用数学、计算数学以及统计学等知识,数学理论和方法是反演问题研究的基础.在科学的发展史上,反演问题现已成为一个最活跃和令人振奋的交叉学科.

反问题的典型特性是其不稳定性^[1]:首先,问题的解可能不存在,给定的数据可能不符合建立数学模型时的要求;其次,解可能不唯一,即使解存在也可能不唯一;其三,问题的解不连续依赖于观测数据的变化,因而带来数值模拟的不稳定性.

通常,我们把求数学物理反问题的稳定近似解的方法称为正则化方法.国外关于反演问题理论和方法的研究起步较早,在 20 世纪 20 年代 Hadamard 在研究线性偏微分方程的 Cauchy 问题时就对反演问题的不稳定性做了陈述和研究^[6].20 世纪 40 年代 Tikhonov 率领他的工作小组开始了反问题的理论研究^[7],到 60 年代中期,在针对重磁解释中的特殊病态问题时,提出了处理不稳定性问题的至今广泛沿用的 Tikhonov 正则化方法.而 Phillips^[8]也在 1962 年独立地提出了正则化方法.随着该方法的提出,对不稳定性问题的解的研究进入了新的阶段.在相当长的一段时间内,人们的注意力集中在该类方法及其各种推广之上,比如说正则化方法的迭代执行、广义逆方法等等.

Tikhonov 正则化方法适用于解病态方程,是一种通用且有效的方法,但对于某些具有特殊结构的算子和当初始数据具有某些特殊信息时,利用正则化方法的标准算法一般很难获得最佳的解.数学工作者在理论上不断完善,针对主要的两个困难:一方面是最佳正则化参数的选取,另一方面是正则子的合理选择展开了深入的理论研究.Engl^[9-13]和 Hansen^[14-18]在正则化方法的理论和算法方面做了大量研究.

我国在反问题方面的研究起步稍晚,20 世纪 80 年代,在大量实际问题的驱动下,已故的中国科学院院士冯康教授^[19]大力倡导反演问题的研究,有关反演理论和方法在相关的领域中陆续开始展开,越来越多的优秀工作者投入到这一领域的研究中,取得了一系列有关反演问题的理论和方法的优秀的科研成果.如张关泉^[20]对波动方程反演问题的理论和算法进行了系统研究,黄光远等^[21]从控制和脉冲谱角度对反演问题进行论述,程晋等^[22-27]在数学物理反问题的理论和数值方法方面展开研究,对 Tikhonov 正则化方法的参数选取方法进行了进一步的研究,提出了基于反问题条件稳定估计的简单法则,给出了误差估计和算法.刘继军等

[28-33]在正则化的理论及应用,如对逆散射问题的唯一性问题的研究等方面都取得了丰富的成果.

反演问题理论和方法的基础性研究固然重要,但如何把经过发展的反演问题的理论和方法更有效应用到地球物理勘探等实际领域中,仍然是一个十分值得关注和研究的课题.近年来,应用方面的研究一直在发展中.如对于地球物理反演中震波图的最优正则化方法^[34,35],对于图像复原的基于FFT的Tikhonov正则化方法及其各种扩展^[36],对于由遥感和地学中地表参数反演的核驱动模型而发展起来的基于Bayesian统计检验的带先验估计的最小二乘方法^[37],QR方法^[38],NTSVD方法^[39]等.

在我国,栾文贵^[40]首先将正则化方法应用于地球物理反演中.近些年来地球物理工作者围绕此方法,进行了广泛研究,应用到各种地球物理反演中,如地震偏移反演^[41]、波阻抗反演^[42]、测井反演^[43]、大地电磁反演^[44]、电磁逆散射^[45]等各个领域,取得了丰富的成果.

特别是,由于计算数学学科中最优化方法的发展,人们越来越多地关注把此类方法与正则化方法相结合,应用到地球物理反演问题的求解上.例如,人们提出了正则化的Broyden方法,正则化的Gauss-Newton方法,正则化的CG方法,截断共轭梯度方法、内点方法、信赖域方法等等^[1].这些方法应用于实际问题,均取得了良好的结果,但这方面的研究还远远没有成熟,值得广大科研工作者共同去完成.

1.3 研究目标与主要内容

本文的目的是将最优化方法和正则化方法相结合,研究发展新的病态矩阵方程解算方法.重点将围绕Tikhonov提出的处理不适定问题的正则化方法,结合Active-Set、差分进化等优化算法,以及选取带双参数的正则化子等方面展开研究.通过反演算法的合理设计和算法的可行性测试,并以双频电磁波电导率成像反演的计算为例,在理论模型的模拟计算分析的基础上,再进行电导率反演成像的实验.以此说明各算法在地球物理反演计算中的实用可能性.

第 2 章 地球物理反演问题

本章首先将介绍地球物理反演的基本概念, 以及地球物理中的一些典型反演问题, 并且通过这些典型事例认识反演问题的不适定性.

2.1 地球物理反演问题

地球物理学中有正(演)问题和反(演)问题. 反问题是相对于正问题而言的. 通常, 由已知的一个地球物理模型去求其响应函数的问题叫做正(演)问题(direct problem), 而由地球物理观测资料去反推地球物理模型的问题则统称为反(演)问题(inverse problem). 与正问题相比, 反问题的发展历史相对较短, 一直到 20 世纪 60 年代的中期, 才成为一个真正的研究领域, 引起应用科学家和数学家的广泛重视和深入研究.

地球物理反演的研究内容包括了地球物理学中的任何一类反问题. 地球物理反演是地球物理学中近年来发展很快的利用地球表面及钻孔中观测到的物理数据、推测地球内部介质物理参数分布和变化的新方法. 通常把地球物理反演概括为: 地球物理学中的反演理论就是研究把地球物理学中的观测数据映射到相应的地球物理模型的理论和方法^[46].

在油气勘探、电磁、信号/图像处理、遥感等众多地球物理领域, 有着各种各样的反问题. 下面我们首先来介绍几个地球物理中的反问题.

例 2.1 地震勘探中的波阻抗反演, 参见【1】.

假设地层为水平层状介质, 考虑如下的一维声波方程

$$\rho(h) \frac{\partial^2 x}{\partial t^2} = \frac{\partial}{\partial h} \left(k(h) \frac{\partial x}{\partial h} \right), \quad (2-1)$$

其中 $x = x(h, t)$ 为波场, 它是时间 t 和深度 h 的可微函数, $\rho(h)$ 和 $k(h)$ 分别表示

介质的密度和体变量模. 如果令 $v(h) = \sqrt{\frac{k(h)}{\rho(h)}}$ 为波传播速度, $\tau(h)$ 表示波传到

深度为 h 时的走时: $\tau(h) = \int_0^h \frac{ds}{v(s)}$, 并采用波阻抗表示式 $\sigma(\tau) = \rho(\tau)v(\tau)$, 则(2-1)

可以表示成

$$\sigma(\tau) \frac{\partial^2 x}{\partial t^2} = \frac{\partial}{\partial \tau} \left(\sigma(\tau) \frac{\partial x}{\partial \tau} \right), \quad (2-2)$$

假定：当 $x \in [a, b]$ 时， $Z(x) = -H + z(x)$ ，当 $x \notin [a, b]$ 时， $Z(x) = -H$ 。

由于重力场的作用，可在地表测得一隅角为

$$\Delta_g = \frac{\partial V}{\partial z} \Big|_{z=0}$$

其中 V 是由于密度差 $\rho = \rho_2 - \rho_1$ 引起的位势。

由于

$$V = \int_s \frac{\rho}{2\pi} \ln \left(\frac{1}{r} \right) d\xi d\eta, \quad r = \sqrt{(x-\xi)^2 + (z-\eta)^2},$$

有

$$\Delta_g = -\frac{\rho}{2\pi} \int_a^b \int_{-H}^{-H+z(\xi)} -\frac{\partial}{\partial \eta} \ln \left(\frac{1}{r} \right) d\xi d\eta \Big|_{z=0} = \frac{\rho}{2\pi} \int_a^b \ln \frac{(x-\xi)^2 + H^2}{(x-\xi)^2 + (H-z(\xi))^2} d\xi,$$

而隅角 Δ_g 是可以测定的。我们的反问题是要确定边界 $z(x)$ ，这就归结为求解如下的非线性第一类 Fredholm 积分方程

$$Az \equiv \int_a^b \ln \frac{(x-\xi)^2 + H^2}{(x-\xi)^2 + (H-z(\xi))^2} d\xi = y(x)$$

其中 $y(x) = \frac{2\pi}{\rho} \Delta_g$ 。这里 A 是一个非线性算子。

地球物理学中反演问题的目的是根据观测数据寻求相应的地球物理模型中的参数。在地球物理学中，尽管问题千差万别，但把观测数据与地球物理模型参数联系起来的数学模型却只有线性和非线性两大类。绝大多数观测数据和模型参数之间虽都不满足线性关系，但是，在一定近似条件下均可简化或者近似简化为线性关系。所以，线性反演一直以来是地球物理学家非常关心的问题。

从以上给出的反问题的例子可以看出，设有一个数学模型描述了某一物理过程，记 x 为该数学模型的未知特性， y 为实验观测结果，而 K 是一个把 x 映像成 y 的算子，该过程可以简单地写成

$$Kx = y.$$

反问题就是由已知 K 和 y 近似地来求未知量 x 。当算子 K 为线性算子时，称其为线性反问题，否则称为非线性的反问题。

下述的图示(图 2-2 和图 2-3)^[1]可以用来简单的描述上述过程。在其中图 2-2 (a) 描述了正问题的发生过程；图 2-2 (b) 和图 2-3 描述了反问题的发生过程。

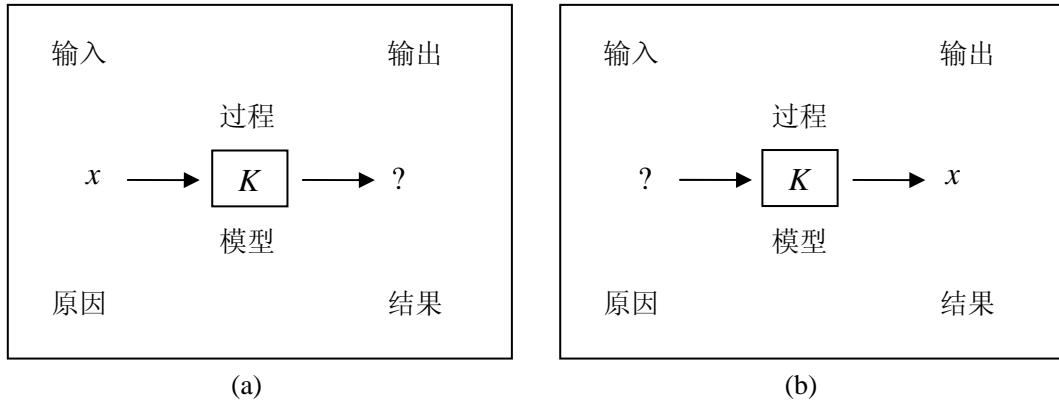


图 2-2 正问题示意图

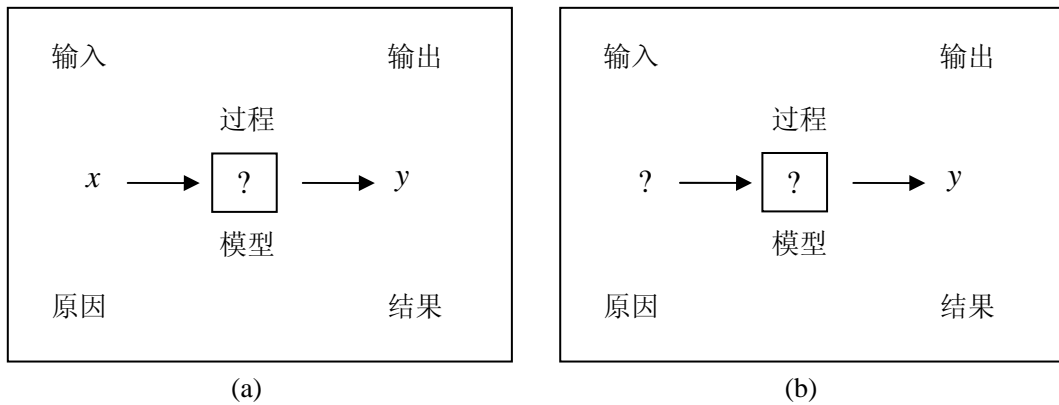


图 2-3 反问题示意图

2.2 反演问题的不适定性

因观测数据的不完备性、模型与算子的近似性等原因，各类地球物理反演问题都存在不程度的不适定性。不适定性的一般定义如下^[1]：

定义 2-1 令 $K: X \rightarrow Y$ 是一个算子， X 和 Y 为相应的空间，如果算子方程

$$Kx = y \quad (2-6)$$

全部满足下述条件：

- (a) 对任意给定的 $y \in Y$ ，都存在 $x \in X$ ，使得(2-6)成立；
- (b) (2-6)式的解是唯一存在的；
- (c) 问题(2-6)的解连续依赖于右端项 y 的变化，即：若 $Kx_0 = z_0$ ， $Kx = z$ ，则当 $z \rightarrow z_0$ 时，有 $x \rightarrow x_0$ 。

则说算子方程是适定的；反之，只要上面三个条件有任何一个不满足，这个问题就是不适定的。

问题的不适定性对其数值求解将产生本质的影响. 直接将算子方程近似求解的标准方法用于不适定方程时往往会产生无意义的数值结果. 接下来我们针对一些不适定问题的例子, 从理论分析和数值求解两个方面来说明不适定问题的特殊性.

例 2.3 第一类积分方程的数值解, 如重力测定中的反问题就转化为第一类积分方程. 参见【31】.

考虑

$$\int_0^1 e^{ts} x(s) ds = y(t) \stackrel{\text{def}}{=} \frac{e^{t+1} - 1}{t+1}, 0 \leq t \leq 1.$$

该方程的唯一精确解是 $x(t) = e^t$. 在对该问题求数值解时, 取步长 $h = \frac{1}{n}$, 用复合梯形公式可以得到左端积分项的近似为

$$\int_0^1 e^{ts} x(s) ds \approx h \left(\frac{1}{2} x(0) + \frac{1}{2} e^t x(1) + \sum_{j=1}^{n-1} e^{jht} x(jh) \right).$$

最后由线性代数方程组

$$h \left(\frac{1}{2} x_0 + \frac{1}{2} e^{ih} x_n + \sum_{j=1}^{n-1} e^{jih^2} x_j \right) = y(ih), i = 0, 1, \dots, n$$

求 $x(jh)$ 的近似值 $x_j (j = 0, 1, \dots, n)$. 下表 2-1 对不同的 n 值, 给出表示数值解和真解在点 $t = jh$ 的误差 $x(jh) - x_j$.

表 2-1 数值解和真解在点 $t = jh$ 的误差 $x(jh) - x_j$

t	$n=4$	$n=8$	$n=16$	$n=32$
0	0.44	-3.08	1.08	-38.21
0.25	-0.67	-38.16	-25.17	50.91
0.5	0.95	-75.44	31.24	-116.45
0.75	-1.02	-22.15	20.03	103.45
1	1.09	-0.16	-4.23	-126.87

从上面表中可以看到, 随着分点 n 的增多, 左端积分项的计算精度应该不断提高, 但是结果却显示, 数值解和真解在每一点之间的误差反而越来越大. 显然, 这是一个无意义的数值结果^[47].

该问题本质上是一个第一类 Fredholm 积分方程的求解问题, 就是所说的典型的不适定问题. 除此以外, 第一类 Volterra 积分方程的求解问题, 同样也是不

适定问题. 关于这种不适定问题, 必须引进正则化的解法, 而不能简单地直接用数值积分的方法来求解. 否则就会产生上面例子出现的这种现象.

其实, 这类问题求解的特殊性很早就引起了物理学家的注意. 早在 1962 年, 美国物理学家 Philips 就考虑了由积分方程

$$\int_{-6}^6 (1 + \cos \frac{\pi(x-y)}{3}) f(x) dx = (6+y)(1 - \frac{1}{2} \cos \frac{\pi y}{3}) - \frac{9}{2\pi} \sin \frac{\pi y}{3}, |y| \leq 6$$

求解 $f(x)$ 的问题. 通过直接离散积分项、再求解在节点处对应的线性代数方程组的办法, 所得到的数值结果同样随着离散节点数目的增加而越来越坏^[48].

例 2.4 有限维线性代数方程组的求解, 参见【31】.

对 $n \times n$ 阶实方阵 A 和 n 维实向量 u 及 z , 考虑线性代数方程组

$$Az = u$$

的解 z .

当 $\det A \neq 0$ 时, 对任意给定的 $u \in R^n$, 该方程组存在唯一解 $z = A^{-1}u \in R^n$, 并且解连续依赖于右端项.

当 $\det A = 0$ 时, 该方程并不能保证对任意的右端项 u 都有解. 而当对某个右端项 u_0 有解时, 解一定是不唯一的. 因此, 此时该问题的求解是不适定的.

这个简单例子表明了解的存在性、唯一性和连续依赖性不一定是完全独立的. 也就是引起问题不适定性的三个原因在某些条件下是有联系的, 换言之, 对本问题而言, 解的适定性等价于 $Az = 0$ 只有平凡解 $z = 0$. 也就是说, 对本问题, 只要有解的唯一性, 就有了解的存在性和连续依赖性, 即解的适定性.

上面两个不同的具体例子有一个共同的特点^[31], 就是在原问题具有不适定性时, 如果应用经典的方法去求问题的解时, 会出现这样的结果: 要么是给定的输入数据不一定能保证解的存在性或唯一性; 要么是输入数据的微小变化会引起相应解的巨大变化, 而且这种变化已经使得用通常方法求得的对解变得毫无意义.

不适定性本质上是由于信息不足或者过量, 也就是一些输入数据太多或者太少, 使得待求的解要么不足要么过定造成的. 所以, 要恢复问题的适定性尤其是稳定性, 就需要采用添加信息、改变拓扑度量等方法. 对添加信息, 通常是对待求的解作某些附加假定, 或者在一个较小的集合上求解, 以此得到稳定性. 而对于改变拓扑度量, 只能是从数学理论研究的角度而言, 因为从应用的角度来考虑, 度量方式在给定的具体应用问题中是难以随便改变的. 新的度量(模)选择必须尽可能有合理的物理解释, 数学意义上过于抽象的模对数学方法的具体应用是不利的. 另一方面, 在数值计算的过程中, 过于弱的适定性结果, 由于计算精度的干扰, 也是很难表现出来的, 这个现象已经被许多数值计算结果所证实.

2.3 电导率反演问题

电导率等电磁参数是岩石介质最重要的物性参数之一。以电磁参数分布为调查对象的电磁方法在工程及地质探查、环境及水文地质调查、油气储层勘探开发等方面占有重要地位。电磁方法的探测深度与分辨能力是互为矛盾的。探测深度通常是必须满足的硬条件，因此数据处理方法便成为提高分辨率的关键。除具规则几何形体的简单电性结构外，很难用解析表达式将观测数据和介质电磁参数分布联系起来，唯一可行的是采用数值离散反演。无论采用何种近似进行线性化，都将形成一个规模巨大、高度奇异的矩阵方程。

这里考虑非均匀介质双频电磁波电导率成像问题。在一阶近似模型条件下，具有如下形式

$$E_1 - E_2 = 20 \log_{10}(\sqrt{f_1 / f_2}) + 8.6859 \sqrt{\pi \mu_0} \cdot (\sqrt{f_2} - \sqrt{f_1}) \cdot \left(\sum_{j=1}^n \sqrt{\sigma_j} \cdot r_j \right)$$

其中 σ_j 是单元 j 的电导率， r_j 为第 j 个单元中射线段的长度(如图2-4)。

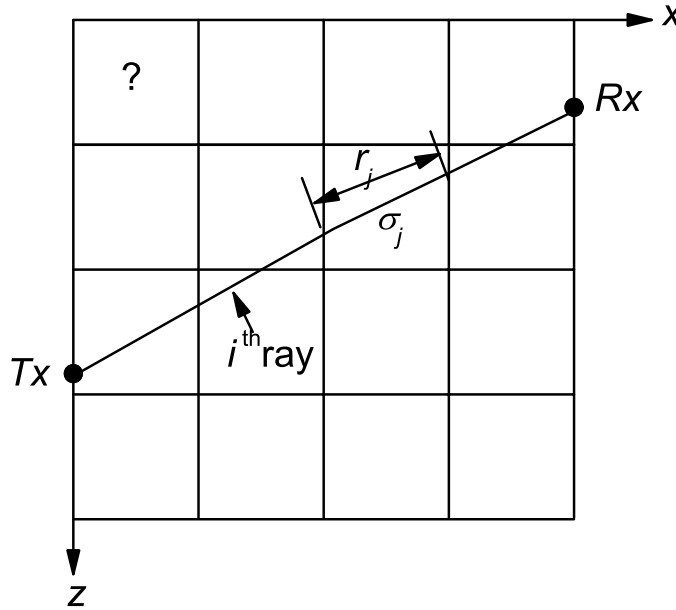


图 2-4 井间地层剖分和接收路径示意图

记 $C_1 = 57.944 / (\sqrt{f_2} - \sqrt{f_1})$, $C_2 = 20 \log_{10}(\sqrt{f_1 / f_2}) \cdot C_1$, $d = C_1 \cdot (E_1 - E_2) - C_2$, 则

$$\sum_{j=1}^n \sigma_j^{1/2} r_j = d. \quad (2-7)$$

对 m 个观测数据，由(2-7)式可得双频电磁波电导率成像方程

$$A_{l \times n} m = d + e, \quad (2-8)$$

其中 $m = (\sigma_1^{1/2}, \sigma_2^{1/2}, \dots, \sigma_n^{1/2})^T$ 为未知的电导率参数，系数矩阵 $A = \{a_{ij}\}$ 为贡献度矩阵，其元素 a_{ij} 的物理意义是射线 i 在单元 j 内的实际长度， d 为由观测数据得到的已知向量， e 为误差向量，包含测量误差和线性化误差。求解方程组(2-8)，即可得到电导率成像图。根据方程(2-8)进行反演成像，避开了单频电磁衰减成像时发射天线的初始辐射场强度与辐射方向因子的计算，有助于提高重建图像的精度。

但是这里的方程组(2-8)是一个高度病态的矩阵方程，应用经典的数值求解方法去求解时，会出现毫无意义的结果。

第 3 章 正则化方法概述

通常，我们把求数学物理反问题的稳定近似解的方法称为正则化方法。自从 20 世纪 60 年代 Tikhonov^[5]提出求解不适定问题的正则化理论和方法以来，人们一直致力于此类方法的理论研究及其对各种实际问题的应用推广。地球物理工作者也将此方法应用到地球物理的各个领域，取得了丰富的成果。Tikhonov 正则化方法适用于解病态方程，是一种通用且有效的方法，但对于某些具有特殊结构的算子和具有某些特殊信息的初始数据，利用正则化方法的标准算法一般很难获得最佳的解。这里主要有两个困难，一个是最佳正则化参数的选取，另一个是正则子的合理选择。这也是正则化方法的两个核心问题。

本章我们首先介绍反演问题的数值解法，然后对正则化方法的数学理论基础做一个较详细的说明，最后回顾近年的一些应用到地球物理反演计算的正则化算法的研究进展。

3.1 反演问题的数值解法

地球物理反演目的主要是根据观测数据求取相应的物理模型，从而能够获得解，并根据解来得到有关真实模型的地球物理信息^[1]。

假设 X 是参数空间， Y 是观测空间。在地球物理科学研究中，总是需要求解下列的观测数据方程

$$(k_i, x) = y_i, \quad i = 1, 2, \dots, M, \quad (3-1)$$

其中 $x \in X, y \in Y, k_i$ 为核函数。设 k_i 定义在 $L^2(\Omega)$ 上，则(3-1)可以写为如下的积分方程

$$\int k_i(t)x(t)dt = y_i, \quad i = 1, 2, \dots, M. \quad (3-2)$$

而(3-2)正好是第一类 Fredholm 方程

$$\int k(s, t)x(t)dt = y(s), \quad i = 1, 2, \dots, M \quad (3-3)$$

的特例。很显然，矩方程(3-2)可以表述成为抽象的算子方程

$$Kx = y. \quad (3-4)$$

因为这里 $y \in Y$ ，而 Y 是观测空间，因此自然会带有各种误差干扰，也称为噪音干扰，这里用 e 表示各种误差，当假设误差是具有可加性的，那么上面的方程可以写成

$$Kx + e = y^e. \quad (3-5)$$

为了获得上述问题的解，一个很自然的想法就是令

$$\|e\| \rightarrow \min. \quad (3-6)$$

这样就把这个问题的求解变成了下列变分问题的求解:

$$\arg \min_x J_1[x, y^e] \stackrel{\text{def}}{=} \frac{1}{2} \|Kx - y^e\|^2. \quad (3-7)$$

在前一节我们谈到反演问题通常具有不适定性, 而问题(3-2)由于算子是紧的, 所以是一个不适定的逆问题. 由于 K^{-1} 有可能不存在, 或者这个问题本身没有解, 还有可能 x 不连续依赖于 y 的变化, 因此如果直接利用 $x = K^{-1}y$ 进行求解是不可取的. 这里必须利用正则化的技巧, 即转变为求解下列的变分问题:

$$\arg \min_x J_2[x, y^e] \stackrel{\text{def}}{=} \frac{1}{2} \|Kx - y^e\|^2 + \frac{\alpha}{2} \Omega[x]. \quad (3-8)$$

下面我们就来看转变为不同的变分问题时, 分别采用的一些常用的计算方法. 通常, 我们将数值求解方法分为直接方法和迭代方法.

3.1.1 直接方法

直接方法主要包括我们常见的广义逆方法、SVD 方法、TSVD 方法、以及 Tikhonov 正则化方法的直接求解等.

3.1.1.1 截断奇异值分解方法

利用广义逆反演方法求解

$$K_{m \times n} x_{n \times 1} = y_{m \times 1} + e_{m \times 1}$$

得到未知参数 x 的最小范数最小二乘解为:

$$x_{LS} = K^+ y, \quad (3-9)$$

其中, K^+ 是 K 的 Moore-Penrose 广义逆. 设 $\text{rank} K = n$, K 的奇异值分解式为

$$K = U \Sigma V^T = \sum_{i=1}^n \sigma_i u_i v_i^T, \quad (3-10)$$

其中, 正交矩阵 $U = (u_1, \dots, u_n)$, $V = (v_1, \dots, v_n)$, 而

$$\Sigma = \begin{bmatrix} \Sigma_n & 0 \\ 0 & 0 \end{bmatrix},$$

其中 $\Sigma_n = \text{diag}\{\sigma_1, \dots, \sigma_n\}$, $\sigma_1 \geq \dots \geq \sigma_n > 0$ 是 K 的奇异值. 相应地, K^+ 的奇异值分解式为

$$K^+ = V \Sigma U^T = \sum_{i=1}^n \sigma_i^{-1} v_i u_i^T. \quad (3-11)$$

将(3-11)式代入(3-9)式, 得到最小范数最小二乘解的奇异值分解式为

$$x_{LS} = K^+ y = \sum_{i=1}^n \frac{u_i^T y}{\sigma_i} v_i,$$

在第 N 步对其进行截断, 得到截断的奇异值分解式为

$$x_N = \sum_{i=1}^N \frac{u_i^T y}{\sigma_i} v_i, \quad (3-12)$$

可以看出, 当 $N = n$ 时, $x_N = x_{LS}$. 但是通过截断, 适当去除 $(n - N)$ 个大误差项, 恢复了一些解的主要特性, 同时也丧失了一些解的精确性^[1,49].

3.1.1.2 Tikhonov正则化方法

给定参数 $\alpha > 0$, 选择 $x_\alpha \in X$ 使得

$$\min \left\{ \|Kx - y\|^2 + \alpha \|x - x_0\|^2 \right\}, \quad (3-13)$$

其中 α 为 Tikhonov 正则化参数, x_0 为参数 x 的一个先验值, (3-13)式的极小解 $x_\alpha \in X$ 是连续依赖于参数 $\alpha > 0$ 的. 在实施正则化方法过程中, 确定合适的正则化参数是核心问题之一, 参数的选取应使得观测数据的误差与实施正则化过程产生的误差相一致. 考虑 Tikhonov 正则解 $x_\alpha = (\alpha I + K^* K)^{-1} K^* y, \alpha > 0$. 应用 K 的奇异系统, 它可以表示为^[1]:

$$x_\alpha = \sum_{i=1}^n \frac{\sigma_i}{\alpha^2 + \sigma_i^2} u_i^T y v_i.$$

3.1.2 迭代方法

变分问题(3-8)的求解还可以用迭代法来求解. 典型的方法有如最速下降方法、非线性最小二乘法、Landweber 迭代法、共轭梯度(简称 CG)法、Newton 型方法等.

3.1.2.1 LSQR方法

LSQR 方法是利用 Lanczos 法求解最小二乘问题, 是 Paige 和 Sanders^[50]在 1982 年提出的. 与 SVD 法相比, LSQR 方法具有计算量小的优点, 并且能很容易地利用矩阵的稀疏性简化计算, 因而适合求解大型稀疏问题.

LSQR 方法具体步骤如下(详见【50, 51】):

(1) 初始化

$$\begin{aligned} \beta_1 u_1 &= b, \quad \alpha_1 v_1 = K^T u_1, \quad w_1 = v_1, \\ x_0 &= 0, \quad \rho_1 = \alpha_1, \quad \phi_1 = \beta_1. \end{aligned} \quad (3-14)$$

(2) 双对角化

$$\begin{aligned}\beta_{i+1}u_{i+1} &= Kv_i - \alpha_i u_i, \\ \alpha_{i+1}v_{i+1} &= K^T u_{i+1} - \beta_{i+1}v_i.\end{aligned}\quad (3-15)$$

(3) 正交变换

$$\begin{aligned}\rho_i &= \sqrt{\bar{\rho}_i^2 + \beta_{i+1}^2}, \quad c_i = \bar{\rho}_i / \rho_i, \quad s_i = \beta_{i+1} / \rho_i, \\ \theta_{i+1} &= s_i \alpha_{i+1}, \quad \bar{\rho}_{i+1} = -c_i \alpha_{i+1}, \quad \varphi_i = c_i \bar{\varphi}_i, \quad \bar{\varphi}_{i+1} = s_i \bar{\varphi}_i.\end{aligned}\quad (3-16)$$

(4) 更新

$$\begin{aligned}x_{i+1} &= x_i + (\varphi_i / \rho_i)w_i, \\ w_{i+1} &= v_{i+1} - (\theta_{i+1} / \rho_i)w_i.\end{aligned}\quad (3-17)$$

在上述迭代公式中，选取非负常数 α_i, β_i 单位化相应的右端向量，如 $\beta_1 u_1 = b$ 表示 $\beta_1 = \|b\|$, $u_1 = b/\|b\|$.

3.1.2.2 共轭梯度(CG)法

定义函数

$$f(x) = \|Kx - y\|^2 = (Kx - y, Kx - y), \quad x \in X \quad (3-18)$$

则 $\nabla f(x) = 2K^*(Kx - y) \in X$. 如果 $(Kp, Kq) = 0$, 我们称两元素 p, q 关于算子 K 共轭. 共轭梯度法的算法可描述如下^[52]:

①初始化: $x_0, d_0 = y - Kx_0, p_1 = s_0 = K^*d_0$, 其中 x_0 为先验值;

②对 $k=1, 2, \dots$, 若 $s_{k-1} \neq 0$, 则计算

$$\begin{aligned}q_k &= Kp_k; \\ \alpha_k &= \|s_{k-1}\|^2 / \|q_k\|^2; \\ x_k &= x_{k-1} + \alpha_k p_k; \\ d_k &= d_{k-1} - \alpha_k q_k; \\ s_k &= K^*d_k; \\ \beta_k &= \|s_k\|^2 / \|s_{k-1}\|^2; \\ p_{k+1} &= s_k + \beta_k p_k;\end{aligned}$$

直到 $s_{k-1} = 0$.

由共轭梯度法产生的点列 $\{x_k\}$ 满足

$$\|y - Kx_k\| \rightarrow \min \{\|y - Kx\|\}. \quad (3-19)$$

3.2 正则化方法的数学基础

上节已经提到，正则化方法主要是为了解决第一类算子方程特别是第一类积分方程而发展起来的，解决不适定性的典型方法是变分正则化方法.

正则化方法解决不适定问题的关键是正则化参数的选取和收敛速度的估计. 求解不适定问题的正则化算子的构造和相应的函数空间(解空间、数据空间)有着密切的联系. 在一般的度量空间上, Tikhonov 引进了变分正则化方法. 在性质更好的空间, 如 Banach 空间、Hilbert 空间上, 相关的结果也是成立的. 其基本思想是引入所谓的“稳定泛函”. 下面简单介绍在 Hilbert 空间的框架下, 线性不适定问题正则化方法的有关结果. 以下各节主要内容取自文献【31, 52】及【1】.

3.2.1 一般正则化方法

假定 K 是由 Hilbert 空间 X 到 Hilbert 空间 Y 的一个紧的线性算子, 则由 $Kx = y \in Y$ 求解 $x \in X$ 的问题是不适定的. 很多的反问题都可以化为这类算子方程的求解问题, 而正则化理论求解这类不适定问题的基本思想是对解加上先验的条件, 以恢复问题的适定性, 主要是唯一性和稳定性. 对任何实际观测到的精确的 y , 方程

$$Kx = y \quad (3-20)$$

存在唯一解 x . 这意味着 $K: X \rightarrow K(X)$ 是一一对应的. 在什么情况下能保证 $K: X \rightarrow Y$ 是一一对应的(这意味着对任意给定的 $y \in Y$, $Kx = y$ 都有唯一解), 或者说, 如何选取 Y 使得 $K(X) = Y$, 这是反问题研究中的一个重要理论问题, 即解的存在唯一性. 这里主要讨论反问题研究中的另一个重要理论问题, 即不适定问题解的稳定性.

设有误差的观测数据 y^δ 满足

$$\|y^\delta - y\| \leq \delta.$$

如何由 y^δ 求 x 的近似值 x^δ ? 一般说来, x^δ 不能由对应的方程

$$Kx^\delta = y^\delta$$

直接来求, 原因是首先: 当 $y^\delta \notin K(X)$ 时, 该方程无解; 其次: 即使 $y^\delta \in K(X)$, 由于 K^{-1} 是无界的, 当 $\delta \rightarrow 0$ 时, 也不能保证由上式得到的解 $x^\delta \rightarrow x$.

为了使得在 $y^\delta \notin K(X)$ 的时候, 也能用合适的方法构造出近似解 x^δ , 并保证 x^δ 对 y^δ 的连续依赖性, 必须构造 $K^{-1}: K(X) \rightarrow X$ 的有界近似算子 $R: Y \rightarrow X$.

定义 3-1 一族有界线性算子 $R_\alpha: Y \rightarrow X$ ($\alpha > 0$) 称为是式 $Kx = y$ 的正则化解算子, 如果对任意给定的 $x \in X$, 满足

$$\lim_{\alpha \rightarrow 0} R_\alpha Kx = x$$

成立, 其中 α 称为正则化参数.

通过这个定义, 我们可以看出:

(1) 对于任意一个固定的 $\alpha > 0$, R_α 是有界的;

(2) 在 X 上, $\alpha \rightarrow 0$ 时, $R_\alpha K$ 逐点收敛于单位算子 I ;

(3) 定义等价于: 对任意给定的 $y \in K(X)$, $R_\alpha y \rightarrow K^{-1}y$ 成立;

(4) 如果 K 是一般的紧算子, 也可以定义其连续的正则化算子 $R_\alpha: Y \rightarrow X$ ($\alpha > 0$), 使其去逼近不连续算子 K^{-1} , 但 R_α 不一定是线性的. 此时, R_α 的有界性要求需要改为 $R_\alpha y$ 对 $y \in Y$ 的连续性.

下面来看正则化算子的性质.

定理 3-1 设 $\dim X = \infty$, $K: X \rightarrow Y$ 是一个紧算子, 而 R_α 是 K^{-1} 的正则化算子, 则

(1) R_α 关于 α 不是一致有界的. 即有序列 $\{\alpha_j\} \rightarrow 0$ 使得 $j \rightarrow \infty$ 时 $\|R_{\alpha_j}\| \rightarrow \infty$.

(2) $\alpha \rightarrow 0$ 时 $\|R_\alpha K - I\| \rightarrow 0$ 不成立.

根据正则化算子的定义, 对右端的精确数据 $y = Kx$, 正则化解 $R_\alpha y$ 当 $\alpha \rightarrow 0$ 时当然收敛于精确解 x . 下面讨论右端数据不精确的情况.

令 $y \in K(X)$ 是右端的精确数据, 而 $y^\delta \in Y$ 是满足误差式 $\|y^\delta - y\| \leq \delta$ 的观测数据. 记

$$x^{\alpha, \delta} = R_\alpha y^\delta$$

是由扰动数据 y^δ 构造的 $Kx = y$ 的精确解 x 的近似值. 由估计

$$\begin{aligned} \|x^{\alpha, \delta} - x\| &\leq \|R_\alpha y^\delta - R_\alpha y\| + \|R_\alpha y - x\| \\ &\leq \|R_\alpha\| \|y^\delta - y\| + \|R_\alpha Kx - x\| \\ &\leq \delta \|R_\alpha\| + \|R_\alpha Kx - x\| \end{aligned}$$

可以看出误差分成两部分: 第一项是由输入数据的误差 $\delta > 0$ 产生的解的误差, 但它被正则化算子的模 $\|R_\alpha\|$ 放大了. 第二项表示在精确右端数据 y 处由正则化算子 R_α 逼近不连续算子 K^{-1} 产生的误差 $\|(R_\alpha - K^{-1})y\|$, 当 $\alpha \rightarrow 0$ 时它趋于 0. 由定理 3-1, 对任意给定的 $\delta > 0$, 当 $\alpha \rightarrow 0$ 时 $\delta \|R_\alpha\| \rightarrow \infty$. 因此, 正则化参数 α 的选取必须保持某种平衡. 一方面, 由近似解对输入数据的误差 $\delta > 0$ 的稳定性要求需要 $\|R_\alpha\|$ 很小(即 α 不能太小); 另一方面, 正则化算子 R_α 对不连续算子 K^{-1} 的逼近性要求 $\|(R_\alpha - K^{-1})y\|$ 很小(即 α 越小越好). 这样求解不适定方程的正则化方法的基本问题是: 决定一种策略 $\alpha = \alpha(\delta)$ 使得

$$\delta \|R_\alpha\| + \|R_\alpha Kx - x\|$$

极小化, 且当 $\delta \rightarrow 0$ 时有 $\delta \|R_\alpha\| + \|R_\alpha Kx - x\| \rightarrow 0$.

这就是正则化方法的一个核心问题——正则化参数的选取.

解决该问题的基本思想是用 α 分别来估计 $\|R_\alpha\|$ 和 $\|R_\alpha Kx - x\|$ 的误差界

$\delta c_1(\alpha) + c_2(\alpha)$ ，再极小化函数 $\delta c_1(\alpha) + c_2(\alpha)$ 以确定 $\alpha = \alpha(\delta)$ (需用到精确解的先验估计)，并且该极小化函数在 $\delta \rightarrow 0$ 时也趋于 0。

定义 3-2 正则化参数的取法 $\alpha = \alpha(\delta)$ 称为是允许的，如果在 $\delta \rightarrow 0$ 时，同时成立

$$\alpha(\delta) \rightarrow 0, \sup_{x \in X} \left\{ \|R_{\alpha(\delta)} y^\delta - x\| : \|Kx - y^\delta\| \leq \delta \right\} \rightarrow 0.$$

正则化参数的取法有先验(a-priori)和后验(a-posteriori)两种方式。先验取法是基于精确解的光滑性条件，而后验取法是基于数据误差水平信息和误差数据本身。在实际中，由于精确解的光滑性条件很难预先给出，所以通常正则化参数的取法是采用后验方式。

下面我们来看允许的 $\alpha = \alpha(\delta)$ 的取法。首先看一个相关的定理。

定理 3-2 (Picard 定理) 设 $K: X \rightarrow Y$ 是线性紧算子， K 的奇异系统为 $(\mu_j(>0), x_j, y_j)(j=1, 2, \dots)$ 。方程(3-20)可解的充分必要条件是

$$y \in N(K^*)^\perp, \sum_{j=1}^{\infty} \frac{1}{\mu_j^2} |(y, y_j)|^2 < \infty,$$

且在方程(3-20)为可解时，其解的表达式为

$$x = \sum_{j=1}^{\infty} \frac{1}{\mu_j} (y, y_j) x_j.$$

这里， X, Y 是 Hilbert 空间， $K: X \rightarrow Y$ 是紧算子， K^* 是 K 的伴随算子。 $N(K^*)$ 为 K^* 的零空间， $N(K^*)^\perp$ 为其正交。 $(\mu_j(>0), x_j, y_j)(j=1, 2, \dots)$ 是 K 的奇异系统，即成立

$$Kx_j = \mu_j y_j, \quad K^* y_j = \mu_j x_j, \quad j=1, 2, \dots.$$

一方面 $\mu_j \rightarrow 0$ ，另一方面由 Picard 定理，当 $y \in K(X)$ 时，

$$x = \sum_{j=1}^{\infty} \frac{1}{\mu_j} (y, y_j) x_j,$$

因此构造正则化算子 R_α 的方法本质上就是找到一种方法把算子 K 的小奇异值 μ_j 过滤掉。下面的结果表明了这种一般构造的可能性。

定理 3-3 设线性紧算子 K 的奇异系统为 $(\mu_j(>0), x_j, y_j)(j=1, 2, \dots)$ ，函数

$$q(\alpha, \mu): (0, \infty) \times (0, \|K\|) \rightarrow \mathbf{R}$$

满足下列性质：

- (1) 对一切的 $\alpha > 0$ 和 $0 < \mu < \|K\|$ ，成立 $|q(\alpha, \mu)| \leq 1$ ；
- (2) 存在函数 $c(\alpha)$ 使得对一切的 $0 < \mu < \|K\|$ 成立 $|q(\alpha, \mu)| \leq c(\alpha)\mu$ ；
- (3a) 对每一个给定的 $\mu(0 < \mu < \|K\|)$ 成立 $\lim_{\alpha \rightarrow 0} q(\alpha, \mu) = 1$ ，

则有下列结论：

(1) 如下定义的算子 $R_\alpha: Y \rightarrow X (\alpha > 0)$:

$$R_\alpha y = \sum_{j=1}^{\infty} \frac{q(\alpha, \mu_j)}{\mu_j} (y, y_j) x_j, \quad y \in Y,$$

是一个正则化算子, 且有估计 $\|R_\alpha\| \leq c(\alpha)$;

(2) 如果取 $\alpha = \alpha(\delta)$ 在 $\delta \rightarrow 0$ 时满足 $\alpha(\delta) \rightarrow 0$ 及 $\delta c(\alpha(\delta)) \rightarrow 0$, 则 $\alpha = \alpha(\delta)$ 是允许的取法.

在这个定理中, $R_\alpha: Y \rightarrow X (\alpha > 0)$ 是正则化算子意味者对一切的 $x \in X$ 当 $\alpha \rightarrow 0$ 时 $R_\alpha Kx \rightarrow x$ 或者 $\|R_\alpha Kx - x\| \rightarrow 0$. 由于精确解 x 是未知的, 能否用 α 来估计 $\|R_\alpha Kx - x\|$? 由于 $\|R_\alpha K - I\| \rightarrow 0$ 一般不成立 (详见文献【31】), 该估计需要对 x 的一些更严格的假定. 在求正则化解 $x^{\alpha, \delta}$ 时, 由于

$$\|x^{\alpha, \delta} - x\| \leq \delta \|R_\alpha\| + \|R_\alpha Kx - x\| \leq \delta c(\alpha) + \|R_\alpha Kx - x\|,$$

故对允许的取法 $\alpha = \alpha(\delta)$, 即可用 δ 来估计近似解的误差.

在对精确解 x 的一些先验假定和比条件(3a)的更强的条件下, 可解决该问题.

定理 3-4 设 $q(\alpha, \mu)$ 满足定理 3-3 中的(1)和(2). 且满足

(A1) 如果 $x = K^* z \in K^*(Y)$, 并且满足

$$(3b) \text{ 对一切的 } \alpha > 0 \text{ 和 } 0 < \mu \ll \|K\|, \text{ 成立 } |q(\alpha, \mu)| \leq c_1 \frac{\sqrt{\alpha}}{\mu},$$

则有估计

$$\|R_\alpha Kx - x\| \leq c_1 \|z\| \sqrt{\alpha}.$$

(A2) 如果 $x = K^* Kz \in K^* K(X)$, 并且满足

$$(3c) \text{ 对一切的 } \alpha > 0 \text{ 和 } 0 < \mu \ll \|K\|, \text{ 成立 } |q(\alpha, \mu)| \leq c_2 \frac{\sqrt{\alpha}}{\mu^2},$$

则有估计

$$\|R_\alpha Kx - x\| \leq c_2 \|z\| \alpha.$$

与定理 3-3 相比较, 此定理的优点在于: 如果知道精确解 x 的一些先验信息, 则可以选取适当的过滤函数 $q(\alpha, \mu)$, 使得我们可以估计出正则化解趋向于精确解的速度. 但是, 该定理中所要求的精确解的先验信息一般是很难验证的.

满足定理 3-4 中条件的 $q(\alpha, \mu)$ 很多, 对应于不同的 $q(\alpha, \mu)$, 可得到不同的 $\alpha = \alpha(\delta)$ 的允许取法以及相应的误差估计.

例如, 下面同时列出满足条件(1)、(2)、(3b)及(3c)的三个函数 $q(\alpha, \mu)$:

$$(1) \quad q(\alpha, \mu) = \frac{\mu^2}{\alpha + \mu^2} \text{ 对应于 } c(\alpha) = \frac{1}{2\sqrt{\alpha}}, c_1 = \frac{1}{2}, c_2 = 1;$$

(2) $q(\alpha, \mu) = 1 - (1 - \alpha\mu^2)^{1/\alpha}, 0 < \alpha < \frac{1}{\|K\|^2}$, 对应于

$$c(\alpha) = \sqrt{\frac{a}{\alpha}}, c_1 = \frac{1}{\sqrt{2a}}, c_2 = \frac{1}{a};$$

(3) $q(\alpha, \mu) = \begin{cases} 1, & \mu^2 \geq \alpha \\ 0, & \mu^2 < \alpha \end{cases}$, 对应于 $c(\alpha) = \frac{1}{\sqrt{\alpha}}, c_1 = 1, c_2 = 1$.

这三种取法所得的正则化解的精度本质上是一致的, 因为它们对应于 R_α 的相同的估计 $\|R_\alpha\| \leq 1/\sqrt{\alpha}$, 只是表现形式不同. 函数 $q(\alpha, \mu)$ 的前两种取法可以避免 K 的奇异值而构造正则化方法, 而 $q(\alpha, \mu)$ 的第三种取法称为谱截断方法, 在工程上应用很广泛. 谱截断对应的正则化解是

$$x^{\alpha, \delta} = \sum_{\mu_j^2 \geq \alpha} \frac{1}{\mu_j} (y^\delta, y_j) x_j.$$

由定理 3-4, 可以得到对 $\|x^{\alpha, \delta} - x\|$ 的估计.

设 $x = K^* z \in K^*(Y), \|z\|_Y \leq E$. 如取 $\alpha(\delta) = c \frac{\delta}{E} (c > 0)$, 则有

$$\|x^{\alpha, \delta} - x\| \leq \left(\frac{1}{\sqrt{c}} + \sqrt{c} \right) \sqrt{\delta E};$$

设 $x = K^* K z \in K^* K(Y), \|z\|_X \leq E$. 如取 $\alpha(\delta) = c \left(\frac{\delta}{E} \right)^{\frac{2}{3}} (c > 0)$, 则有

$$\|x^{\alpha, \delta} - x\| \leq \left(\frac{1}{\sqrt{c}} + \sqrt{c} \right) E^{\frac{1}{3}} \delta^{\frac{2}{3}}.$$

用过滤掉算子的小奇异值的方法来求紧算子逆的正则化算子, 已有很长的历史, 它对理论研究是很方便的. 但是在具体问题的应用上, 应构造避开 K 的奇异值的正则化方法, 因为算子的奇异值通常是很难求的. $q(\alpha, \mu)$ 的前两种取法可以完成这种构造.

接下来将介绍 Tikhonov 正则化方法.

3.2.2 Tikhonov 正则化方法

在有限维空间近似求解超定的线性代数方程组 $Kx = y$ 时, 所用的方法是求最小二乘解, 即在有限维空间 X 上极小化连续泛函 $\|Kx - y\|$. 该问题一定是可解的. 但是如果 X 是无限维而 K 是紧的, 则该极小化问题是不适定的.

定理 3-5 设 X, Y 是 Hilbert 空间, $K: X \rightarrow Y$ 是有界线性算子. 对 $y \in Y$, 存在 $\hat{x} \in X$ 使得

$$\|K\hat{x} - y\| \leq \|Kx - y\|$$

对一切 $x \in X$ 成立的充分必要条件是 \hat{x} 满足

$$K^* K \hat{x} = K^* y,$$

其中 $K^*: Y \rightarrow X$ 是 K 的伴随算子.

一般说来, 不能保证上述方程解的存在唯一性, 因此该极小化问题是一个不适定的问题. 但如果对极小元 \hat{x} 加上进一步的限制(例如是最小模解), 就可以保证极小元的存在唯一性.

用 Tikhonov 正则化方法求解不适定问题的基本想法是在目标函数 $\|Kx - y\|$ 上加上适当的惩罚项, 使得从优化理论的角度, 求新的目标函数的极小元的问题适应, 或者从积分方程理论的角度, 使得极小元满足的方程是一个第二类的方程. 具体说来, 将问题的提法改为: 对有界线性算子 $K: X \rightarrow Y$ 和任给的 $y \in Y$, 求 $x^\alpha \in X$ 使其在 $x \in X$ 上极小化 Tikhonov 泛函

$$J_\alpha(x) = \|Kx - y\|_Y^2 + \alpha \|x\|_X^2,$$

其中 $\alpha > 0$ 称为正则化参数. 此时有

定理 3-6 设 X, Y 是 Hilbert 空间, $K: X \rightarrow Y$ 是有界线性算子, 则

- (1) $J_\alpha(x)$ 在 X 上存在唯一的极小元 x^α ;
- (2) $x^\alpha \in X$ 满足

$$\alpha x^\alpha + K^* K x^\alpha = K^* y.$$

对有界线性算子 $K: X \rightarrow Y$, 还可以进一步证明: 由定理 3-6 确定的最小元 x^α 关于 α 是无限次可微的, 且其导数可递推得到.

定理 3-7 设 X, Y 是 Hilbert 空间, $K: X \rightarrow Y$ 是有界线性算子. 对任意给定的 $\alpha > 0$, x^α 关于 α 是无限次可微的, 且其 n 阶导数 $\omega := \frac{d^n}{d\alpha^n} x^\alpha \in X$ 可由

$$\alpha \omega + K^* K \omega = -n \frac{d^{n-1}}{d\alpha^{n-1}} x^\alpha, n = 1, 2, \dots,$$

或者其等价的变分形式:

$$(K\omega, Kg)_Y + \alpha(\omega, g)_X = -n \left(\frac{d^{n-1}}{d\alpha^{n-1}} x^\alpha, g \right)_X, \forall g \in X$$

递推确定.

借助于泛函 $J_\alpha(x)$ 的极小元 x^α 关于的 α 可微性的上述结果, 可以得到由相容性原理确定正则化参数 α 的一个迭代办法. 在这个办法中, 给定的观测数据 y^δ 的误差水平 δ 可以是未知的, 且可以由观测数据 y^δ 本身来估计出 δ 的界. 其基本出发点是构造极小化函数 $J(x^\alpha)$ 的一个近似的有解析表达式的模型函数, 再由此模型函数近似确定正则化参数.

求 $Kx = y$ 的正则化解还有另一种方法, 即可以通过求一个泛函在全空间上的

极小元，或者直接解一个第二类的方程来确定正则化解 x^α 。它对应于过滤函数

$$q(\alpha, \mu) = \frac{\mu^2}{\alpha + \mu^2},$$

但是避开了 K 的奇异值 μ_j 。再由对应于上述过滤函数的正则化解的理论结果，就可以得到 Tikhonov 正则化方法应用于有误差的观测数据时的收敛性结果。

定理 3-8 设 $K: X \rightarrow Y$ 是紧的线性算子， $\alpha > 0$ 。

(1) $(\alpha I + K^*K)$ 是有界可逆的， $R_\alpha \stackrel{\text{def}}{=} (\alpha I + K^*K)^{-1} K^*: Y \rightarrow X$ 是 $Kx = y$ 的一个正则化解算子， $\|R_\alpha\| \leq \frac{1}{2\sqrt{\alpha}}$ 。对应于近似的右端数据 y^δ ， $Kx = y$ 的 Tikhonov 正则化解 $x^{\alpha, \delta} = R_\alpha y^\delta$ 由

$$\alpha x^{\alpha, \delta} + K^* K x^{\alpha, \delta} = K^* y^\delta$$

唯一确定。正则化参数 $\alpha = \alpha(\delta)$ 只要在 $\delta \rightarrow 0$ 时满足

$$\alpha(\delta) \rightarrow 0 \text{ 及 } \frac{\delta^2}{\alpha(\delta)} \rightarrow 0,$$

就是允许的取法。

(2) 设 $x = K^* z \in K^*(Y)$ ， $\|z\| \leq E$ 。则取 $\alpha(\delta) = c\delta/E$ 时，有估计

$$\|x^{\alpha(\delta), \delta} - x\| \leq \frac{1}{2} \left(\frac{1}{\sqrt{c}} + \sqrt{c} \right) \sqrt{E\delta}.$$

(3) 设 $x = K^* K z \in K^* K(X)$ ， $\|z\| \leq E$ ，则取 $\alpha(\delta) = c(\delta/E)^{2/3}$ 时，有估计

$$\|x^{\alpha(\delta), \delta} - x\| \leq \left(\frac{1}{2\sqrt{c}} + c \right) E^{1/3} \delta^{2/3}.$$

在结果(3)中，其收敛性估计是 $\delta^{2/3}$ 。可以证明，对一般的 Tikhonov 正则化方法，满足 $\|Kx^\delta - y\| \leq \delta$ 的正则化解的误差估计 $\|x^\delta - x\|$ 最好也只能达到 $\delta^{2/3}$ 。因此在精确解的先验条件 $x \in K^* K(X)$ 下，Tikhonov 正则化方法能够得到最优的近似解，其误差是 $\delta^{2/3}$ 。这就是下面的定理 3-9。

定理 3-9 设 $K: X \rightarrow Y$ 是一对一的紧线性算子，且 $K(X)$ 是无限维的。对 $x \in X$ ，如果存在 $\alpha: [0, \infty) \rightarrow [0, \infty)$ 满足 $\alpha(0) = 0$ ，且使得对任何满足 $\|y^\delta - Kx\| \leq \delta$ 的 $y^\delta \in Y$ ，成立

$$\lim_{\delta \rightarrow 0} \|x^{\alpha(\delta), \delta} - x\| \delta^{-2/3} = 0,$$

其中 $x^{\alpha(\delta), \delta} \in X$ 是将式 $\alpha x^\alpha + K^* K x^\alpha = K^* y$ 中 y 换为 y^δ 的解，则 $x = 0$ 。

定理 3-9 的结果表明，如果 Tikhonov 正则化解收敛于精确解 x 的速度比 $\delta^{2/3}$ 还快，则精确解 x 必为 0。

在文献【13,53】中给出了定理 3-8 和定理 3-9 的更一般的形式。

定理 3-10 若精确解满足光滑性条件 $x \in \text{Range}((K^*K)^\nu)$, $0 < \nu \leq 1$, 则当正则化参数的选取满足 $\alpha = c\delta^{\frac{2\nu}{2\nu+1}}$ 时, 就是可取的方法, 并且解有收敛性估计 $\|x^{\alpha(\delta), \delta} - x\| = O(\delta^{\frac{2\nu}{2\nu+1}})$, 其最优收敛速度在 $\nu=1$ 获得, 而且是不可再改进的.

在前述由 Tikhonov 正则化泛函

$$J_\alpha(x) \stackrel{\text{def}}{=} \|Kx - y\|^2 + \alpha \|x\|^2, x \in X, \alpha > 0$$

来构造 $Kx = y$ 的正则化近似解 $x^\alpha = R_\alpha y$ 时, x^α 是由

$$\alpha x^\alpha + K^* K x^\alpha = K^* y$$

唯一确定的. 在由 y 的观测数据 y^δ 来求正则化近似解时, $\alpha = \alpha(\delta)$ 的取法也已给定, 使得在 $\delta \rightarrow 0$ 时 $x^{\alpha(\delta), \delta} \rightarrow K^{-1}y$. 这里 $\alpha = \alpha(\delta)$ 是先验选取的, 即在求解 $\alpha x^\alpha + K^* K x^\alpha = K^* y$ 时必需先给定 $\alpha(\delta)$. 为了讨论 $\alpha = \alpha(\delta)$ 的后验取法, 必须进一步讨论式 $\alpha x^\alpha + K^* K x^\alpha = K^* y$ 的解 x^α 对 α 及 y 的连续依赖性.

仍假定 X 和 Y 是 Hilbert 空间, $K: X \rightarrow Y$ 是一对一的有界线性算子, 并设 $K(X)$ 在 Y 中稠密.

定理 3-11 对 $y \in Y, \alpha > 0$, 记 x^α 是式 $\alpha x^\alpha + K^* K x^\alpha = K^* y$ 的唯一解, 则

- (1) x^α 连续依赖于 α 及 y ;
- (2) 映射 $\alpha \mapsto \|x^\alpha\|_X$ 是单调非增的, 且 $\lim_{\alpha \rightarrow +\infty} x^\alpha = 0$;
- (3) 映射 $\alpha \mapsto \|Kx^\alpha - y\|_Y$ 是单调非减的, 且 $\lim_{\alpha \rightarrow 0} Kx^\alpha = y$;
- (4) 如果 $K^* y \neq 0$, 则 $\alpha \mapsto \|x^\alpha\|_X$ 是严格减的, 而 $\alpha \mapsto \|Kx^\alpha - y\|_Y$ 是严格增的.

这里, Tikhonov 正则化的提法是在整个空间 X 上极小化泛函 $J_\alpha(x)$. 而由定理 3-8 中的(2)及(3)知, 如将精确解 x 限制到一个小的空间中时, 可得到更好的收敛性结果. 因此可在 Tikhonov 正则化方法中考虑更强的惩罚项, 即考虑

$$\|Kx - y\|_Y^2 + \alpha \|x\|_1^2$$

在 $X_1 \subset X$ 上的极小化问题, 其中 $\|\cdot\|_1$ 是 X_1 上的模 (半模). 这是 Tikhonov 正则化方法的一般思想.

考虑正则化方程 $\alpha x^\alpha + K^* K x^\alpha = K^* y$ 在 $\alpha=0$ 时的情况. 如果方程 $K^* v = 0$ 只有零解 $v=0$, 则求解 $Kx = y$ 和 $K^* K x = K^* y$ 是等价的. 此时显然有 $(v, y)_Y = 0$. 这是 Fredholm 选择定理的一个特例: $Kx = y$ 有解的充要条件是 $(v, y)_Y = 0$, 其中 v 满足 $K^* v = 0$. 事实上,

$$(v, y)_Y = (v, Kx)_Y = (K^* v, x)_X = 0, \forall v \in N(K^*).$$

下面再给出 Tikhonov 正则化方程式 $\alpha x^\alpha + K^* K x^\alpha = K^* y$ 的一个等价形式, 据

此可得到求解 $\alpha x^\alpha + K^* K x^\alpha = K^* y$ 的一个迭代算法, 并给出对 $\alpha x^\alpha + K^* K x^\alpha = K^* y$ 的进一步解释.

定理 3-12 正则化方程 $\alpha x^\alpha + K^* K x^\alpha = K^* y$ 等价于弱耦合的方程组

$$\begin{cases} Kx^\alpha - \sqrt{\alpha} v^\alpha = y, \\ K^* v^\alpha + \sqrt{\alpha} x^\alpha = 0, \end{cases}$$

其中 v^α 称为伴随变量.

事实上, 从上方程组中消去 v^α 即得式 $\alpha x^\alpha + K^* K x^\alpha = K^* y$. 另一方面, 把式 $\alpha x^\alpha + K^* K x^\alpha = K^* y$ 写为 $K^*(Kx^\alpha - y) + \alpha x^\alpha = 0$, 并记 $\sqrt{\alpha} v^\alpha = Kx^\alpha - y$, 即得上述的方程组.

对待定的非零实数 β , 上方程组可改写为

$$\begin{cases} Kx^\alpha - \sqrt{\alpha} v^\alpha = y, \\ \sqrt{\alpha} \beta K^* v^\alpha + \alpha \beta x^\alpha - x^\alpha + x^\alpha = 0, \end{cases}$$

据此可给出原方程组的迭代解法为 (β 依赖于 n):

$$\begin{cases} \sqrt{\alpha} v_n^\alpha = Kx_n^\alpha - y, \\ x_{n+1}^\alpha = x_n^\alpha - \beta_n [\alpha x_n^\alpha + K^* \sqrt{\alpha} v_n^\alpha]. \end{cases}$$

该迭代程序通过求解两个正问题来得到序列 x_{n+1}^α : 给定 x_n^α , 先计算原正问题 Kx_n^α 产生 v_n^α , 再计算伴随正问题 $K^* v_n^\alpha$ 产生 x_{n+1}^α . 注意到 $J_\alpha(x)$ 的梯度为 $2(\alpha x^\alpha + K^*(\sqrt{\alpha} v^\alpha))$, 因此上述的第二步迭代就是求泛函极小值的共轭梯度法, 从而可以由此确定常数 β_n . 因此上述迭代格式就是求解 $Kx = y$ 的正则化的伴随共轭梯度方法.

Tikhonov 正则化方法把正则化泛函

$$J_\alpha(x) = \|Kx - y\|^2 + \alpha \|x\|^2, x \in X, \alpha > 0$$

的极小元 x^α 作为方程(3-20)的正则化近似解. 该正则化解一方面使得 $\|Kx^\alpha - y\|$ 较小(从而 $Kx = y$ 的近似解), 另一方面通过罚项 $\alpha \|x\|^2$ 来约束解的大小或者解的振荡性, 保证解的稳定性(从而正则化解). 从优化理论的角度来看, 该二次优化问题可以看成是下述两个带约束的优化问题的惩罚函数方法:

- (1) 对给定的 $\rho > 0$, 在对解的约束条件 $\|x\| \leq \rho$ 下极小化偏差函数 $\|Kx - y\|$;
- (2) 对给定的 $\delta > 0$, 在对偏差的约束条件 $\|Kx - y\| \leq \delta$ 下极小化解的模 $\|x\|$.

Tikhonov 正则化方法最早是对线性不适定的问题 $Kx = y$ 的求解提出的. 在某些情况下, **Tikhonov 正则化方法**可以改变原问题的某些特性和结构. 另一方面, **Tikhonov 正则化方法**中引进了正则化项来减弱原不适定问题近似解的振荡性, 即要求所求的正则化解 x^α 具有一定的光滑性, 用它来近似原问题的解 x . 显然, 如

果原问题的精确解 x 本身就不光滑甚至具有一些奇性, 近似解 x^α 就不可能完全反应这种奇性. 换句话说, Tikhonov 正则化方法得到稳定近似解的代价是模糊了原问题真实解的奇性, 而这一奇性特征在一些问题的求解中是非常重要的, 甚至正是所要求的. 这是 Tikhonov 正则化方法的另一个缺点, 即解的过度光滑化.

基于 Tikhonov 正则化方法存在的上述问题, 人们引进了所谓的迭代正则化方法, 典型的就 Landweber-Fridman 正则化方法. 详见文献【54】和【55】.

3.2.3 正则化参数选择方法

正则化参数 α 的选取对所求解的性态起着关键的作用. 如果 α 太小, 则对问题的谱的改善没有起到什么作用, 即解的不稳定性仍然存在; 如果 α 太大, 所得到的新问题是稳定地求解了, 但该问题已经与原问题相去甚远, 是一个相当糟糕的逼近, 所求的根本不是原来问题的解了. 最优的正则化参数选取应当兼顾这两种情况.

正则化参数的选择, 始终是一个重要而具有魅力的研究课题^[13-16,56-63], 通常, 对参数 α 的选取有先验的和后验的两类策略. 前者, 在求出正则解以前就已经将正则参数确定下来, 而且可以是多值的, 于是自然地提出是否存在“最优正则化参数”的问题. 先验选取正则化参数 α , 主要依据经验就可以了, 通常都取 $\alpha \in (0,1)$. 从另一角度来看, 正则解与准确解之间的误差显然与正则参数的选取有关, 这也导致了使解的误差极小化的正则参数的存在性与确定方法的研究. 不少先验的策略具有理论分析的价值, 但是在实际中常常难以验证其赖以适用的条件. 因此, 关于确定正则化参数的后验策略, 也就是在计算正则解的过程中, 根据一定的原则来确定与原始数据的误差水平相匹配的正则化参数的研究很盛行, 其方法也很实用. 其基本要求是: 对于待求解施加定性的或定量的信息, 以便使正则参数与原始数据资料的误差水平相匹配. 关于正则化参数的确定问题, 王彦飞在文献【1, 52】中做了详尽的总结:

I. Tikhonov 先验估计

如果算子 $K: X \rightarrow Y$ 是精确给定的, 设 X, Y 是 Hilbert 空间, 求解算子方程

$$Kx = y \quad (3-21)$$

的 Tikhonov 正则化方法指的是求解 Euler 方程

$$x^\alpha = (K^*K + \alpha D)^{-1} K^* y,$$

其中 D 为人为取定的约化算子, 要求是正半定算子, 比如可取为扩散算子. 在数据 y 扰动为 y^δ 的情形, 可以求解相应的 Euler 方程

$$x^{\alpha,\delta} = (K^*K + \alpha D)^{-1} K^* y^\delta.$$

Tikhonov 的先验估计指的是适当选取正则化参数 α , 可以使得 $x^{\alpha,\delta}$ 趋于问题的真实解.

前面已经讨论了 Tikhonov 正则化方法的一般思想，并且在对精确解的某种先验假定条件下，给出了正则化解对精确解的收敛速度。例如，定理 3-8 在精确解 x 属于 K^* 或 K^*K 的值域的条件下，给出了正则化参数 $\alpha = \alpha(\delta)$ 的取法及相应的正则化解 $x^{\alpha(\delta),\delta}$ 的收敛速度。然而，对很多具体的问题，去验证精确解的这类先验条件是非常困难的，甚至是不可能的。文献【22】介绍了正则化参数的另一种选取方法，它建立在原问题的某种条件稳定性的基础上，而不是直接要求精确解的某种先验条件。

II. Morozov 偏差原理

在实际问题中，表征原始资料 y^δ 中的误差水平参数 δ 在不少情况下是可以获取或近似得到的，在这种情况下，一种广为采用的后验策略是所谓的偏差原理。它最早是由 Morozov^[64]于 1966 年提出。

对 y 的有误差的右端数据 y^δ ，由于 $\|y^\delta - y\| \leq \delta$ ，即输入数据的精度只是 δ ，要求偏差 $\|Kx - y^\delta\|$ 过分小于输入数据误差 δ 是没有意义的。因此一个自然的想法就是在满足 $\|Kx - y^\delta\| \leq \delta$ 的函数集上确定 x 使得 $\|x\|$ 尽可能地小。

设 $\|y^\delta - y\| \leq \delta$ ，欲求解算子方程(3-21)，则方程的合适的解应当满足

$$x \in X, \quad \|Kx - y^\delta\| \leq \delta.$$

如果 $\text{Range}(K)$ 非闭，即使 $\text{Null}(K) = \{0\}$ ，集合

$$S = \{x : x \in X, \|Kx - y^\delta\| \leq \delta\}$$

也是无界的，因此应当寻求一个最小模的最小二乘解：

$$\begin{aligned} & \|x\| \rightarrow \min \\ & \text{s.t. } \|Kx - y^\delta\| \leq \delta. \end{aligned} \tag{3-22}$$

注意到集合 S 为闭凸集，因此(3-22)式的解在边界达到，即可记为

$$\begin{aligned} & \|x\|^2 \rightarrow \min \\ & \text{s.t. } \|Kx - y^\delta\|^2 = \delta^2. \end{aligned}$$

引入 Lagrange 乘子 λ ，上述等式约束问题等价于下面的无约束优化问题

$$\|x\|^2 + \lambda \|Kx - y^\delta\|^2 \rightarrow \min,$$

并且 λ 满足偏差方程

$$\|Kx^{\alpha,\delta} - y^\delta\|^2 = \delta^2,$$

其中 $\alpha = \frac{1}{\lambda}$ 。

在原始资料 y^δ 中的误差水平参数 δ 可以获取或近似得到的情况下，Morozov 的偏差原理是十分有效的一种正则化参数选择方法。很多比较有意义的结果都是根据该原理获得的^[56,64,65]有关应用偏差原理决定正则参数及求解反问题的数值算法(如 Newdon 法、快速收敛法、重开始的共轭梯度法等)，可以参考相关文献

【52, 53, 56, 57】.

理论上, Morozov 相容性原理用

$$\|Kx - y^\delta\|^2 = \delta^2$$

来确定正则化参数 $\alpha(\delta)$. 由于 $\alpha(\delta) \rightarrow 0$, 不失一般性, 我们总假定 $0 < \alpha < 1$.

有时由 Morozov 相容性原理的方程 $\|Kx - y^\delta\|^2 = \delta^2$ 确定正则化参数后得到的正则化解仍然不很理想. 一个更一般的推广是由

$$\|Kx - y^\delta\|^2 + \alpha^\gamma \|x^\alpha\|^2 = \delta^2$$

来确定正则化参数 $\alpha(\delta)$, 其中 $\gamma \in [1, +\infty]$ 是一个给定的参数. 该方法称为吸收 Morozov 相容性原理(damped Morozov principle)^[64]. 当 $\alpha \in (0, 1)$, $\gamma = +\infty$ 时式 $\|Kx - y^\delta\|^2 + \alpha^\gamma \|x^\alpha\|^2 = \delta^2$ 即为 $\|Kx - y^\delta\|^2 = \delta^2$. 注意, 当联立求 α 时, 既需要 y^δ 也需要 δ . 而且, 当用数值方法直接解方程 $\|Kx - y^\delta\|^2 = \delta^2$ 时, 数值方案并不总是十分有效的. 例如, 当用 Newton 迭代法求解时, 由于函数的导数很小, 会出现迭代解的收敛速度很慢的问题.

III. Engl 误差极小化准则

在 Hilbert 空间的理论框架下, 容易导出在误差水平未知时决定正则化参数的 Engl 的误差极小化准则^[13].

用 $x^{\alpha, \delta}$ 表示误差水平为 δ 时相应于参数 α 的正则解, x^α 为相应于精确右端项 y 的正则解, 则 $\|Kx^{\alpha, \delta} - y^\delta\|$ 表示“残差”. 从逼近的角度来看, 应使 α 越小越好; 然而, 从数值稳定的角度考虑. 则参数 α 取得越大越好.

关于正则解 $x^{\alpha, \delta}$ 与真解 x 的误差, 我们有下述估计:

$$\|x^{\alpha, \delta} - x\| \leq \|x^{\alpha, \delta} - x^\alpha\| + \|x^\alpha - x\| \leq \|R_\alpha\| \|y^\delta - y\| + \|x^\alpha - x\|, \quad (3-23)$$

其中 $R_\alpha = (K^*K + \alpha D)^{-1}$. 注意在(3-23)式中含有未知的 x 和 y , 故还不具有实用性. 但由于 $\lim_{\alpha \rightarrow 0} \|x^\alpha - x\| = 0$, 故存在常数 $c_1 > 0$, 使得 $\|x^\alpha - x\| \leq c_1$. 同时, 又有 $\|R_\alpha\| \leq 1/\sqrt{\alpha}$, 从而, 若令 $y \approx Kx^\delta$, 可得到正则解的残差上界的近似估计:

$$\|x^{\alpha, \delta} - x\| \leq \frac{\|Kx^{\alpha, \delta} - y^\delta\|}{\sqrt{\alpha}} + c_1 \quad (c_1 \text{ 为常数}).$$

类似地, 也应有

$$\|x^{\alpha, \eta} - x\| \leq \frac{\|K_h x^{\alpha, \eta} - y^\delta\|}{\sqrt{\alpha}} + c_2 \quad (c_2 \text{ 为常数}).$$

由于正则解 $x^{\alpha, \delta}$ 或 $x^{\alpha, \eta}$ 对于任何 $\alpha > 0$ 都是数值稳定的, 故应当使相应的残差极小化; 而从数值稳定的角度来看, 参数 α 又要尽可能取大些才好. 于是, 选择使上述误差上界为极小的参数, 即决定这样的 α , 使得

$$\phi(\alpha) = \frac{\|Kx^{\alpha,\delta} - y^\delta\|}{\sqrt{\alpha}} \rightarrow \min$$

或

$$\phi_h(\alpha) = \frac{\|K_h x^{\alpha,\eta} - y^\delta\|}{\sqrt{\alpha}} \rightarrow \min$$

就很自然了，因为它同时兼顾了“好的逼近”与“好的数值稳定性”的要求。

IV. Goncharsky 广义偏差原理

广义偏差原理是由 Tikhonov 的学生 Goncharsky 等人对于 Morozov 的偏差原理进行改造和推广而得到的。

设算子及右端项均近似地已知，即： $\|K - K_h\| \leq h, \|y^\delta - y\| \leq \delta$ ，于是相应的可行解集为

$$C_\eta = \{x : \|K_h x - y^\delta\| \leq \delta + h \|x\|\},$$

其中 $\eta = (\delta, h)$ 。

设 $x \in D \subset X$ ，我们定义算子方程(3-21)的不相容度为

$$\mu_\eta(y^\delta, K_h) = \inf_{x \in D} \|K_h x - y^\delta\|.$$

显然，如果 $y^\delta \in \overline{K_h D}$ ，则 $\mu_\eta(y^\delta, K_h) = 0$ ，其中，“—”表示闭包。

引理 3-1 设 $\|y^\delta - \bar{y}\| \leq \delta$ ，其中 $\bar{y} = Kx$ ， $x \in D$ ，而 $\|K - K_h\| \leq h$ ，则当 $\eta \rightarrow 0$ 时， $\mu_\eta(y^\delta, K_h) \rightarrow 0$ 。

称

$$\rho_\eta^\lambda(\alpha) = \|K_h x^{\alpha,\eta} - y^\delta\|^2 - (\delta + h \|x^{\alpha,\eta}\|)^2 - (\mu_\eta^\lambda(y^\delta, K_h))^2$$

为广义偏差，其中 $\lambda = \lambda(\eta) \xrightarrow{\text{def}} 0$ 。于是广义偏差原理(GDP)可描述为：

如果 $\|y^\delta\|^2 \leq \delta^2 + (\mu_\eta(y^\delta, K_h))^2$ ，则取 $x^{\alpha,\eta} = 0$ 作为方程(3-21)的近似解；否则，求广义偏差的根 $\alpha^* > 0$ ，取 $x_\eta = x^{\alpha^*,\eta}$ 作为方程(3-21)的近似解。

事实上，即使 $y^\delta \notin \overline{K_h D}$ ，也可以置 $\mu_\eta(y^\delta, K_h) = 0$ 。在可行集 C_η 上求方程(3-21)的最小模的最小二乘解，注意到 C_η 的凸性，不难证明必存在 $x^{\alpha,\eta}$ 满足方程

$$\rho_\eta(\alpha) = \|K_h x^{\alpha,\eta} - y^\delta\|^2 - (\delta + h \|x^{\alpha,\eta}\|)^2 = 0 \quad (3-24)$$

亦称(3-24)为广义偏差方程。

文献【66】提出了一种修改的广义偏差原理(MGDP)。它是基于 $\mu_\eta^\lambda(y^\delta, K_h)$ 的特性。对于 $\gamma \in [0, \infty)$ ，定义

$$\rho_\eta(\alpha) = \|K_h x^{\alpha,\eta} - y^\delta\|^2 - (\delta + h \|x^{\alpha,\eta}\|)^2 - \alpha^\gamma \|x^{\alpha,\eta}\|^2 = 0$$

为新的偏差方程。其中 $\lambda = \lambda(\eta) \rightarrow 0$ ， $\eta = (\delta, h)$ 。数值试验表明 MGDP 的收敛性和稳定性均很好。

V. Arcangeli 准则

Arcangeli 于 1966 年几乎与 Moronov 同时独立地提出了确定正则化参数的一个偏差方法^[67]，他主张由公式

$$\|Kx^{\alpha,\delta} - y^\delta\| - \frac{\delta}{\sqrt{\alpha}} = 0 \quad (3-25)$$

来确定正则化参数。注意到：对于每个固定的 $\delta > 0$ ，函数

$$\rho(\alpha) = \sqrt{\alpha} \|Kx^{\alpha,\delta} - y^\delta\|$$

对 α 连续且单调递增，且有

$$\lim_{\alpha \rightarrow 0} \rho(\alpha) = 0, \quad \lim_{\alpha \rightarrow \infty} \rho(\alpha) = \infty,$$

故存在唯一的一个 $\alpha = \alpha(\delta)$ 满足方程(3-25)。

在一般情况下，用上述几个方法确定正则化参数时，正则解的精度均为 $O(\sqrt{\delta})$ 量级。

VI. 广义 Arcangeli 准则

对于算子和右端项均带有扰动的情形，H.W.Engl^[9]指出， α 应满足下述方程：

$$\rho_\eta(\alpha) = \|K_h^* K_h x^{\alpha,\eta} - K_h^* y^\delta\| = (\delta^p + h^r) \alpha^{-q},$$

此处 $p, r, q > 0$ ， $\eta = (\delta, h)$ ，且假定 $K^* y \neq 0, K_h^* y^\delta \neq 0$ 。可以证明：存在唯一的 $\alpha = \alpha(\eta)$ 满足上述方程。并且通过适当地选取参数 p, r, q ，可使正则解精度进一步提高，例如可能达到 $O(\delta^{2/3})$ 或 $O(\delta^{4/5})$ 的量级^[68]。

用前述的方法确定正则化参数时，都需要预先对于原始数据的误差水平 $\eta = (\delta, h)$ 做出估计。这在有些情况下这是难以做到的，例如在经济领域就是如此^[52]。或者即使勉强做出了估计也是不可靠的^[13]。现在来看在误差水平未知情况下的各种后验准则。这时，下述三个准则是很有用的。

VII. Tikhonov 拟最优准则

Tikhonov 在文献【69】中指出，当数据误差水平 δ 或 η 未知时，可根据下面的拟最优准则：

$$\alpha_{opt} = \min_{\alpha > 0} \left\{ \left\| \alpha \frac{dx^\alpha}{d\alpha} \right\| \right\}$$

来确定正则化参数。其基本思想是：让正则参数 α 以及正则解对该参数的变化率同时稳定在尽可能小的水平上。

记 $\rho_q(\alpha) = \left\| \alpha \frac{dx^\alpha}{d\alpha} \right\|^2, \alpha > 0$ ，则 $\rho_q(\alpha)$ 易由公式

$$\alpha \frac{dx^\alpha}{d\alpha} = -\alpha(K^* K + \alpha I)^{-1} x^\alpha$$

算得。注意在有限维情形总有 $\rho_q(0) = 0$ ，因此在实际计算时应当将初始值 α_0 取

得稍大些.

VIII. 广义交叉校验 (GCV) 准则

广义交叉校验 (GCV) 准则^[70]通常是对算子方程(3-21)的离散形式给出的. 仍将方程的离散化形式记为

$$Kx = y,$$

其中 $K = (a_{ij})_{m \times n}$, $y = (y_1, y_2, \dots, y_m)^T$.

令

$$V(\alpha) = \frac{\|(I - K(\alpha))y^\delta\|^2 / m}{[\text{Tr}(I - K(\alpha))]^2 / m},$$

其中 $K(\alpha) = K_h(K_h^* K_h + \alpha I)^{-1} K_h^*$, $\text{Tr}(I - K(\alpha)) = \sum_{k=1}^m (1 - a_{kk}(\alpha))$, $a_{kk}(\alpha)$ 为 $K(\alpha)$ 的对角元素. 这样可取 α^* 满足

$$V(\alpha^*) = \min V(\alpha).$$

此法源于统计估计理论中的选择最佳 PRESS 准则, 但比它更为稳健.

下面给出方便实现上述准则的一些形式. 注意到恒等式(由熟知的 Sherman-Morrison-Woodburg 定理)

$$\alpha(KK^T + \alpha I)^{-1} = I - K(K^T K + \alpha I)^{-1} K^T,$$

于是 GCV 准则变为极小化如下的函数:

$$V(\alpha) = \frac{\|y - Kx^\alpha\|}{\text{Tr}[(I - K(K^T K + \alpha I)^{-1} K^T)]} = \frac{\|(KK^T + \alpha I)^{-1} y\|}{\text{Tr}[(KK^T + \alpha I)^{-1}]}.$$

Eldén^[71]提出一种求迹 $\text{Tr}[(I - K(K^T K + \alpha I)^{-1} K^T)]$ 的有效方式:

$$\text{Tr}[(I - K(K^T K + \alpha I)^{-1} K^T)] = m - n + \alpha \text{Tr}[(T_\alpha^T T_\alpha)^{-1}],$$

而后者可以递推导出. 其中 T_α 具有形式

$$T_\alpha = \begin{bmatrix} t_{11} & t_{12} & & & 0 \\ & t_{22} & t_{23} & & \\ & & \ddots & \ddots & \\ & & & t_{n-1,n-1} & t_{n-1,n} \\ 0 & & & & t_m \end{bmatrix}$$

记 s_i 为 T_α^{-1} 第 i 行的 l_2 -模, $i = 1, 2, \dots, n$. 则 s_i^2 可以如下计算:

$$s_n^2 = 1/t_m^2,$$

$$s_i^2 = (1 + t_{i,i+1}^2 s_{i+1}^2) / t_{ii}^2, \quad i = n-1, n-2, \dots, 1.$$

于是

$$\text{Tr}[T_\alpha^T T_\alpha)^{-1}] = s_1^2 + s_2^2 + \dots + s_n^2.$$

IX. Hansen 的 L-曲线准则

L-曲线准则是指以 $\log\text{-}\log$ 尺度来描述 $\|x^\alpha\|$ 与 $\|Kx^\alpha - y\|$ 的曲线对比, 进而根据该对比结果来确定正则化参数的方法. 其名称由来是基于上述尺度作图时将出现一个明显的 L-曲线, 见图 3-1.

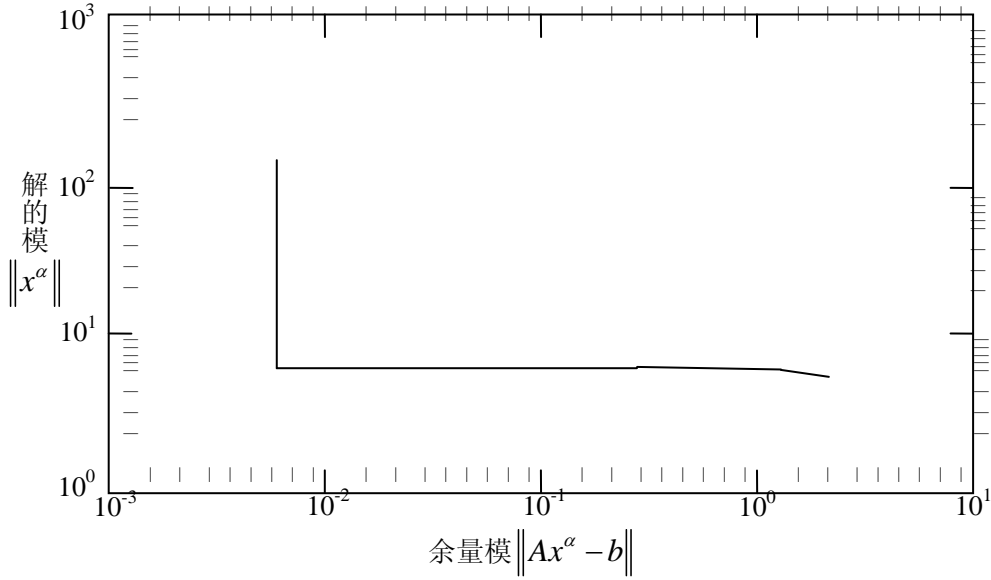


图 3-1 L-曲线示意图

运用 L-曲线准则的关键是给出 L-曲线隅角是数学定义, 进而应用该准则选取参数 α .

Hansen^[16]等建议定义 L-曲线的隅角为 L-曲线在 $\log\text{-}\log$ 尺度下的最大曲率. 令 $\log \|Kx^\alpha - y\|, \theta = \log \|x^\alpha\|$, 则该曲率作为参数 α 的函数为:

$$c(\alpha) = \frac{\rho' \theta'' - \rho'' \theta'}{((\rho')^2 + (\theta')^2)^{3/2}},$$

其中 “'” 表示关于参数 α 的微分.

Engl 等^[9]指出: 在相当多的情况下, L-曲线准则可通过极小化泛函 $\phi(\alpha) = \|x^\alpha\| \|Kx^\alpha - y\|$ 来实现, 即选取 α^* 使得

$$\alpha^* = \arg \{ \inf_{\alpha > 0} \phi(\alpha) \}.$$

这一准则更便于在数值上加以实施. 其实, 我们也可以从另一个角度来诠释该准则的含义: 对于任何给定的 $\alpha > 0$, 正则解 x^α 是数值稳定的, 因而极小化残差 $\|Kx^\alpha - y\|$ 是合理的. 此外, 由于我们要求的是具有极小模的最小二乘解, 故使 $\|x^\alpha\|$ 极小化亦在情理之中. 综上, 极小化乘积 $\|x^\alpha\| \|Kx^\alpha - y\|$ 便可同时达到上述目的.

迄今为止,人们尚未获得关于 L-曲线准则的收敛性结果,事实上有人已经举反例指出了 L-曲线准则的不收敛性^[62],但数值结果表明 L-曲线准则与 GCV 准则一样,有很强的适用性^[16].

3.3 正则化方法在地球物理反演中的应用

20 世纪 60 年代初,地球物理反演在地球物理学家的头脑中扎下了根.反演试图由给定的一组地球物理测量结果重建地下特征.地球物理反演重建可分为两步:观测数据——重建方程;重建方程——解出参数^[72].由观测数据到重建方程所依据的是物理原理与理论,关键是建立观测数据与物理模型参数之间的函数依赖关系.重建方程通常是奇异的或高度病态的大型线性代数方程组.

由重建方程到解出参数本质上是一个纯粹的数学问题.地球物理反演计算中所涉及的大型线性代数方程组通常具有不相容性,不存在精确解,而只能在某种意义上求其近似解.

自从栾文贵^[40]在国内首次将正则化方法应用于地球物理反演以来,近些年来地球物理工作者围绕此方法,进行了广泛研究,应用于各种地球物理的反演计算中,如地震偏移反演、波阻抗反演、测井反演、大地电磁反演、电磁逆散射等各个领域,并取得了丰富的成果.

张宏兵等^[73]将正则化方法应用到波阻抗反演中,通过改变正则参数的数值以及合理地选择正则参数的初值来改善反演结果,提高反演收敛速度.通过理论模型试算和实际资料处理,说明正则参数初值选取对反演进程的影响非常大,应针对不同的势函数和不同的初始波阻抗模型来选取正则参数的初值.如果正则化参数选择过大,只有相对较小的模型扰动量才有可能被接受,使反演结果过于平滑.相反,如果正则化参数选择过小,被接受的可能是相对较大的模型扰动量,使反演结果出现不稳定,发挥不了先验约束作用.随后大量讨论了使用具有边界保护特性的正则化方法以解决波阻抗反演中不适定问题.

范宜仁等^[43]将 Tikhonov 正则化方法应用到地球物理测井反演中.利用变分正则化方法,通过极小泛函构造正则化算子,选取具有最优收敛阶的正则化参数.有效地提高了反演计算的速度和精度,并在实际应用中提高了老井测井资料信息的有效性.

陈小斌等^[44]针对大地电磁正则化反演中正则化因子的选取困难问题提出了自适应正则化反演算法.在该算法中,根据数据目标函数、模型约束目标函数和正则化因子之间的关系,提出了两种正则化因子的自适应调节方法.

熊小芸、聂在平等^[45]也将正则化方法改进用于电磁逆散射问题中.有很多学者在其各自的文献中也对正则化方法进行了介绍.

随着计算数学学科中最优化方法的发展,人们越来越多地关注把此类方法与正则化方法结合,应用到地球物理反演问题的求解上. 王彦飞等^[1,41,66,74-76]成功地应用最优化计算方法和正则化方法研究数学物理反问题,把最优化研究领域中的许多优秀方法应用到求解线性和非线性反演问题上,取得了良好的结果. 他研究了正则化方法及其各种扩展,最早把图像纹理特征提取看成是反演问题,并把正则化方法应用于纹理特征提取上. 特别是,他证明了求解非线性反演问题的信赖域算法也是一种正则化方法,因而可以应用这一方法于反演问题. 同时,他提出了求解数字图像恢复问题的信赖域算法和信号处理超分辨率问题的奇异值分解算法,证明了他们的正则性. 但是有关这些领域的研究还远远没有成熟.

第4章 基于Tikhonov正则化的病态矩阵方程解算

数学家在正则化方法理论上进行了深入的探讨,地球物理学家也逐渐将其应用到实际中.在利用正则化方法处理地球物理反演计算时,通常是采用一般的Tikhonov正则化方法进行计算,也就是采用无约束的拟牛顿法进行优化计算.但针对不同的地球物理问题,可以尝试与不同的优化算法相结合.

本章首先研究将Tikhonov正则化方法与Active-Set算法相结合的计算方法,对此算法的可行性进行详细分析以后,为检验算法解算实际反演问题的实用性,使用该算法对某工区一典型剖面的双频电磁波电导率成像方程^[77,78]进行了重新解算.本章的主要内容已与合作者在地球物理学报等期刊公开发表^[79,80,81].

4.1 反演方法的稳定性

在实际成像问题中,特别是在地球物理成像中,所导出的矩阵方程通常为混定方程.对这类大型的病态矩阵,只能在某种意义上求其近似解.

下面假设各种地球物理反问题在线性化后归结于求解大型稀疏线性方程组的问题,即

$$Am = d, \quad (4-1)$$

式中, A 为 $l \times n$ 阶矩阵,条件数非常大,甚至为奇异矩阵; m 为地球物理参数向量; d 为观测向量.此外,系数矩阵 A 和右端项 d 不可避免地存在误差,包含测量误差和线性化引起的误差.

设 A 的扰动为 δA , d 的扰动为 δd , 相应解的扰动为 δm , 则由(4-1)式可得

$$(A + \delta A)(m + \delta m) = d + \delta d. \quad (4-2)$$

根据数值扰动分析理论^[82], 则有事前误差估计式

$$\frac{\|\delta m\|}{\|m\|} \leq \frac{\text{cond}(A)}{1 - \|A^{-1}\| \cdot \|\delta A\|} \cdot \left\{ \frac{\|\delta A\|}{\|A\|} + \frac{\|\delta d\|}{\|d\|} \right\} \quad (4-3)$$

其中 $\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$ 为矩阵的条件数.当矩阵 A 的条件数非常大,此线性方程组呈现病态时,虽然方程组余量的相对误差已经很小,但解的扰动 δm 的相对误差仍然可能很大.一般的求解方法很难得到方程组(4-1)可靠的近似解,容易造成解的相对误差非常大,偏离真实的解,所以必须采用各种正则化方法进行求解.

已有的解算矩阵方程的算法有多种,如广义逆、LSQR、SVD、TSVD、SASART等等,不同的解法,输出的结果也会不同.而且不同的方法有其适用的不同情况,到目前为止,对具体的大型病态线性方程组还没有尽善尽美的解算方法^[83-89].

4.2 基于 Active-set 算法的正则化解算方法

在一些实际的地球物理反演问题中，常常需要增加先验知识的约束准则，如在电导率成像反演中，需要考虑电导率参数非负的约束条件。而已有的线性反演方法，在考虑物理参数的非负限制方面研究较少。为此，我们在 Tikhonov 正则化思想的基础上对这一情况展开讨论。利用 Active-set 算法^[90]求解最后的非负最小二乘问题，得到电导率成像分布。

考虑到反演参数的非负性，将问题

$$Am = d + e \quad (4-4)$$

转化为如下优化问题：

$$\begin{aligned} \min \quad & \|Am - d\|_2^2 + \lambda^2 \|m - m_{ref}\|_2^2 \\ \text{s.t.} \quad & m \geq 0, \end{aligned} \quad (4-5)$$

其中 A 为 $l \times n$ 阶矩阵， m 为地球物理参数向量， d 为观测向量， e 是由于系数矩阵 A 和右端项 d 而引起的测量误差和线性化误差， m_{ref} 为参数 m 的一个先验值， λ 为 Tikhonov 正则化参数。这里的 m_{ref} 由先验知识得到，本文实际算例中选取 $m_{ref} = (0.1, 0.1, \dots, 0.1)^T$ 。正则化参数 λ 控制着问题(4-5)的解 m_λ 对 A 和 d 的敏感程度，对优化问题的求解起着至关重要的作用。为求解方便，将上述阻尼最小二乘问题通过转变目标函数，变为求解如下非负最小二乘(NNLS)问题：

$$\begin{aligned} \min \quad & \left\| \begin{pmatrix} A \\ \lambda I_n \end{pmatrix} m - \begin{pmatrix} d \\ \lambda m_{ref} \end{pmatrix} \right\|_2^2 \\ \text{s.t.} \quad & m \geq 0. \end{aligned} \quad (4-6)$$

由于矩阵 $\begin{pmatrix} A \\ \lambda I_n \end{pmatrix}$ （其中 I_n 为 n 阶单位阵）为列满秩矩阵，上述 NNLS 问题(4-6)为一个严格凸的优化问题，从而对任意给定的向量 d 和 m_{ref} ，该问题都存在唯一正则化解 m_λ ^[91]。

4.2.1 正则化参数的选取

利用(4-5)求得的正则化解误差包含两个方面：一是由于方程(4-4)右端误差项 e 引起的摄动误差，二是由于方程(4-5)中增加了正则化项后引起的正则化误差。一个好的正则化参数应该合理地平衡所得正则化解的摄动误差和正则化误差。

在实际中，通常无法得知右端误差 $\|e\|_2$ 的大小，故常常采取 L-曲线方法^[15]确定最佳正则化参数。L-曲线方法是目前分析离散病态问题最方便、有效的方法

之一. L-曲线为残差 2 范数 $\|Am-d\|_2$ 和正则化解半范数 $\|m-m_{ref}\|_2$ 之间的关系曲线, 在对数坐标系下, 通常如“L”的形状. L-曲线中垂直部分相应的正则化解的范数 $\|m_\lambda-m_{ref}\|_2$ 对 λ 的选取非常敏感, 而水平部分对应的正则化解的残差 $\|Am_\lambda-d\|_2$ 对 λ 的选取也非常敏感. 所以为了平衡这两个量, 选取参数 λ 对应于 L-曲线中垂直部分和水平部分的交点. 具体如下: 对于连续的正则化参数 λ , 计算曲线

$$(\lg\|Am_\lambda-d\|_2, \lg\|m-m_{ref}\|_2)$$

的曲率, 并找到其曲率极大点, 也就是 L-曲线的角点, 这样就可以确定最佳正则化参数.

在已知 $\|e\|_2$ 大小的情况下, 一般可以采取下面这一种比较简单的方式给出一个最佳的正则化参数 λ_{opt} , 也就是使正则化参数相应的正则化解 $m_{\lambda_{opt}}$ 满足

$$\|Am_{\lambda_{opt}}-d\|_2=\|e\|_2. \quad (4-7)$$

这种方法得到的最佳正则化参数正好对应 L-curve 的角点, 这可从下节实际算例中看出.

4.2.2 Active-set 算法

一般的非负最小二乘问题 (NNLS) 可表示为

$$\begin{aligned} \min_m \quad & \frac{1}{2}\|Am-d\|_2^2, \\ \text{s.t.} \quad & m \geq 0 \end{aligned} \quad (4-8)$$

其中, 这里的 A 为(4-6)式中的列满秩矩阵, $A \in \mathbf{R}^{l \times n}, d \in \mathbf{R}^l, m \in \mathbf{R}^n, l > n$ (其中 $\mathbf{R}^{l \times n}$ 为 $l \times n$ 维向量空间). 此 NNLS 问题可转换为如下二次优化问题:

$$\begin{aligned} \min_m \quad & -(A^T d)^T m + \frac{1}{2} m^T A^T A m \\ \text{s.t.} \quad & m \geq 0. \end{aligned} \quad (4-9)$$

由 KKT 最优化条件可得单调的线性互补问题(LCP)为

$$y = A^T A m - A^T d, \quad y \geq 0, m \geq 0, m^T y = 0. \quad (4-10)$$

如果矩阵 A 列满秩, 则二次优化问题(4-9)为严格凸优化, 线性互补问题(4-10)严格单调, 对任意向量 d 均存在唯一解.

如果满足:

$$\begin{aligned} y &= A^T A m - A^T d, \\ m_i y_i &= 0, \quad i = 1, \dots, n, \end{aligned}$$

则 $(m, y) \in \mathbf{R}^{2n}$ 称为问题(4-10)的互补解.

设 F 和 G 均为 $\{1, \dots, n\}$ 的子集, 且满足 $F \cup G = \{1, \dots, n\}$ 和 $F \cap G = \emptyset$. 相应地, 矩阵 A 的列分块表示为 $A = [A_F, A_G]$, 其中 $A_F \in \mathbf{R}^{m \times |F|}$, $A_G \in \mathbf{R}^{m \times |G|}$, 而 $|F|, |G|$ 分别为矩阵 A_F 和 A_G 的列数. 利用分块矩阵, 可将方程(4-10)写成

$$\begin{bmatrix} y_F \\ y_G \end{bmatrix} = \begin{bmatrix} A_F^T A_F & A_F^T A_G \\ A_G^T A_F & A_G^T A_G \end{bmatrix} \begin{bmatrix} m_F \\ m_G \end{bmatrix} - \begin{bmatrix} A_F^T d \\ A_G^T d \end{bmatrix}, \quad (4-11)$$

其中, $m_F, y_F \in \mathbf{R}^{|F|}$, $m_G, y_G \in \mathbf{R}^{|G|}$, $m = (m_F, m_G)$, $y = (y_F, y_G)$. 在 (4-11) 中, 令 $m_G = 0, y_F = 0$, 可得线性互补问题的一个基本互补解 $((m_F, 0), (0, y_G))$. 具体可由如下两步得到:

第一步, 求解无约束线性最小二乘问题

$$\min_{m_F \in \mathbf{R}^{|F|}} \frac{1}{2} \|A_F m_F - d\|_2^2. \quad (4-12)$$

第二步, 令

$$y_G = A_G^T (A_F m_F - d). \quad (4-13)$$

如果基本互补解 $((m_F, 0), (0, y_G))$ 满足 $m_F \geq 0, y_G \geq 0$, 则称它为可行的, 此时它即为原 NNLS 问题的最优解. 否则, 不可行集 $H = H_1 \cup H_2$ 非空, 其中

$$H_1 = \{k \in F : m_k < 0\}, \quad H_2 = \{k \in G : y_k < 0\}. \quad (4-14)$$

这时, 可按如下方式更新集合 F 和 G :

$$\begin{aligned} F &= F \cup (H_1^* \cup H_2^*), \quad G = G \setminus (H_1^* \cup H_2^*), \\ H_1^* &\subseteq H_1, \quad H_2^* \subseteq H_2, \end{aligned} \quad (4-15)$$

直到找到一个可行的基本互补解为止. 参考文献【92】, 具体算法如下:

Active-set 算法步骤:

- (1) 令 $F = \emptyset, G = \{1, \dots, n\}, m = \mathbf{0}, y = -A^T d$;
- (2) 计算

$$r = \arg \min \{y_i : i \in G\},$$

如果 $y_r < 0$, 令 $H_1^* = \emptyset, H_2^* = \{r\}$, 由(4-15)更新 F 和 G ; 否则, 停止, m 即为原 NNLS 问题的最优解;

- (3) 由(4-11)计算 \bar{m}_F , 如果 $\bar{m}_F \geq 0$, 令 $m = (\bar{m}_F, 0)$, 转第(4)步; 否则, 由

$$\theta = \frac{-m_r}{\bar{m}_r - m_r} = \min \left\{ \frac{-m_i}{\bar{m}_i - m_i} : i \in F \text{ and } \bar{m}_i < 0 \right\}$$

计算出 r , 并令 $m = ((1-\theta)m_F + \theta\bar{m}_F, 0), H_1^* = \{r\}, H_2^* = \emptyset$. 由(4-15)更新 F 和 G , 重做第(3)步.

- (4) 由公式(4-13)计算 y_G , 转第(2)步.

可以证明, 只要矩阵 A 列满秩, Active-set 算法经过有限步可收敛到原 NNLS

问题的最优解^[93].

4.2.3 Active-set算法的数值模拟测试

观察如下一个系数矩阵是 n 阶 Hilbert 矩阵的线性方程组:

$$\begin{bmatrix} 1 & 1/2 & 1/3 & \cdots & 1/n \\ 1/2 & 1/3 & 1/4 & \cdots & 1/(n+1) \\ 1/3 & 1/4 & 1/5 & \cdots & 1/(n+2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1/n & 1/(n+1) & 1/(n+2) & \cdots & 1/(2n-1) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \cdots \\ b_n \end{bmatrix}. \quad (4-16)$$

为了方便地对改进的 Tikhonov 正则化方法与一般正则化方法进行比较, 我们在(4-16)式中取 $n = 25$, 这时上式是一个严重的病态方程组, 其条件数高达 1.2522×10^{19} .

如果假设上式的解为: $x = (1, 1, \dots, 1)^T$, 则可以算出 $n = 25$ 时上式右端项 B 中各元素的值. 下面我们通过已知的 B 和系数矩阵利用基于 Active-set 的 Tikhonov 正则化算法(简称 A-TR 方法)对上式求解, 为了检查它的抗噪性, 这里加入 $SNR=100$ 的高斯白噪声. 这里信噪比 SNR 按如下定义:

$$SNR = 20 \log_{10} \frac{\|d\|_2}{\|d - \tilde{d}\|_2},$$

其计算结果列于下表 4-1.

为定量评价算法的有效性, 分别按如下方式定义相对误差(RE)

$$RE(x^{reg}) = \frac{\|x - x^{reg}\|}{\|x\|} \quad (4-17)$$

其中, x^{reg}, x 为方程组(4-7)的正则化解和精确解.

从与标准解的对比来看, 用改进的 Tikhonov 正则化算法(A-TR 方法)求解上述病态方程组的相对误差很小, 可得到较精确的解. 参看图 4-1 为 A-TR 方法理论模型解的结果.

表 4-1 A-TR 方法计算结果及相对误差

x_i	精确解	A-TR 解结果	绝对误差
1	1	0.99624	0.00376
2	1	1.03772	0.03772
3	1	0.93190	0.06810
4	1	1.00130	0.00130
5	1	1.01628	0.01628
6	1	1.01336	0.01336
7	1	1.01099	0.01099
8	1	1.01086	0.01086
9	1	1.01080	0.01080
10	1	1.00923	0.00923
11	1	1.00579	0.00579
12	1	1.00093	0.00093
13	1	0.99543	0.00457
14	1	0.99009	0.00991
15	1	0.98562	0.01438
16	1	0.98252	0.01748
17	1	0.98114	0.01886
18	1	0.98166	0.01834
19	1	0.98415	0.01585
20	1	0.98857	0.01143
21	1	0.99483	0.00517
22	1	1.00278	0.00278
23	1	1.01228	0.01228
24	1	1.02313	0.02313
25	1	1.03516	0.03516
相对误差(RE)		0.0206	

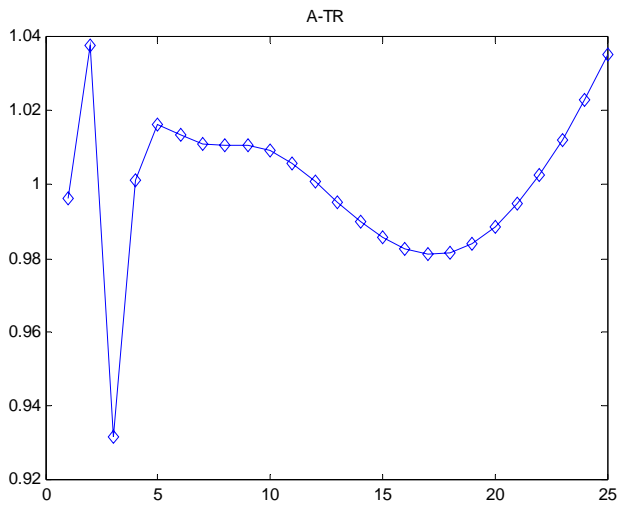


图 4-1 A-TR 方法理论模型解的结果

为了更进一步地研究 A-TR 正则化方法在不同噪声情况下的病态方程组的算法有效性, 增加定义相关系数(CC)

$$CC(x^{reg}) = \frac{\sum_{i=1}^n (x_i - \bar{x})(x_i^{reg} - \bar{x}^{reg})}{\|x - \bar{x}\| \cdot \|x^{reg} - \bar{x}^{reg}\|}, \quad (4-18)$$

其中, x^{reg}, x 为方程组(4-7)的正则化解和精确解, 而 \bar{x}^{reg}, \bar{x} 分别为相对应的正则化解和精确解的均值. 相对误差越小, 相关系数越大, 表示正则化解越接近原方程的真实解.

下面仍然以理论模型 Hilbert 矩阵为例, 在(4-16)式中分别取 $n = 50$ 和 $n = 100$ 进行讨论, 这时其条件数分别为 9.3672×10^{18} 和 1.4690×10^{20} . 这里给定方程组的准确解 $x_i = \cos \frac{i}{2n} \pi$ ($i = 1, 2, \dots, n$), 并由式(4-16)算出准确右端项 b , 再分别加入 SNR=30, SNR=50, SNR=70, SNR=100 的高斯白噪声得到相应的右端项, 然后用 A-TR 正则化方法求解具有误差的病态方程组.

当取方程的阶数 n 分别为 50 和 100 时, 随机进行计算, 得到的一组结果如表 4-2 所示.

从表中也可以看出: 在各种信噪比水平下, A-TR 算法得到的近似解相对误差均比较小, 并且与真解的相关系数也比较大. 当右端项增加的信噪比水平比较高, 也就是右端项的误差相对比较小的情况下, A-TR 算法得到的近似解对真解的逼近程度较高, 几乎接近 1 的值, 而随着信噪比水平的减小, 右端项误差比较大的情况下, A-TR 算法得到的近似解的结果精确度有所降低, 但还是在可以接受的范围.

4.3 基于Active-set算法的电导率反演成像实验

下面对某工区一典型剖面的双频电磁波电导率成像方程进行研究. 现有观测数据 882 个, 未知电导率参数个数为 $20 \times 55 = 1100$ 个. 在这里, 测试方程的系数矩阵 A 为 882×1100 型的, 数值秩为 748, 条件数为 1.6492×10^{31} . 这是一个高度病态的矩阵, 利用通常的方法是很难求解的.

4.3.1 Active-set算法应用实验

下面首先假设测试模型的参数 $m_f = 1$, 得到相应的右端项 $d_f = Am_f$, 然后在理想模型得到的 d_f 上加上一定大小的高斯白噪声得到 \tilde{d}_f , 然后进行反演计算, 检验算法的可行性和可靠性.

这里信噪比 SNR 按如下定义:

$$\text{SNR} = 20 \log_{10} \frac{\|d\|_2}{\|d - \tilde{d}\|_2}.$$

再利用 4.2 节中的所述的算法反演得到 m ，并和假定的 m_f 进行比较。在 SNR 分别为 100、50 的情况下，计算结果分别如图 4-2、图 4-3 所示。

在第一个算例，信噪比 $\text{SNR}=100$ ，引起的右端项误差 $\|e\|_2$ 为 2.88×10^{-4} ，由 L-曲线方法得到的最佳正则化参数 $\lambda_{\text{opt}} = 2.878 \times 10^{-7}$ ，相应的正则化解 $m_{\lambda_{\text{opt}}}$ 在 l_2 范数下的相对误差为 11.45%，残差 $\|Am_{\lambda_{\text{opt}}} - d\|_2 = 2.69 \times 10^{-4}$ 。

在第二个算例中，加入噪音的信噪比减小到 50，右端项误差 $\|e\|_2$ 增至 8.71×10^{-2} ，由 L-曲线方法得到的最佳正则化参数也增至 $\lambda_{\text{opt}} = 5.328 \times 10^{-4}$ ，最后所得的正则化解相对误差随之增大到 14.85%，残差 $\|Am_{\lambda_{\text{opt}}} - d\|_2 = 9.03 \times 10^{-2}$ 。

上面的两个算例说明了由 L-曲线方法确定最佳正则化参数，得到的正则化解近似地满足方程 $\|Am_{\lambda_{\text{opt}}} - d\|_2 = \|e\|_2$ ，也说明了本文算法对这种问题求解的有效性，即使误差比较大，此改进的正则化算法仍能得到比较好的反演结果。

4.3.2 双频电导率成像反演验证

在本节，将根据实际测得的观测数据，反演地层电导率。图 4-4 中 a 图为反演所得的离散电导率数值，b 图为根据文献【78】中所述方法将离散数据插值后得到的电导率图像。在实际计算中，由 L-曲线方法得到的最佳正则化参数 λ 为 21.5。重建图像所反映的电性结构特征与钻探所揭露的剖面地质特征基本一致。在电导率成像图的下部，高导异常体的边界比较清晰，连续性也比较好，与钻探资料和地质规律吻合较好。这些事实表明，用本文算法所得的双频电磁波电导率成像能得到精细可靠的成像图。

表 4-2 不同信噪比水平下 A-TR 正则化解的比较

	信噪比(SNR)	相对误差(RE)	相关系数(CC)
n=50	100	0.0195	0.9990
		0.0169	0.9993
		0.0243	0.9985
	70	0.0562	0.9924
		0.0496	0.9945
		0.0650	0.9883
	50	0.1202	0.9630
		0.1491	0.9462
		0.1186	0.9642
	30	0.2803	0.7830
		0.2655	0.8251
		0.2680	0.8334
n=100	100	0.0468	0.9943
		0.0163	0.9993
		0.0245	0.9985
	70	0.0715	0.9869
		0.0784	0.9841
		0.0546	0.9926
	50	0.1413	0.9486
		0.1404	0.9481
		0.1514	0.9392
	30	0.2799	0.7914
		0.2600	0.8391
		0.3094	0.7254

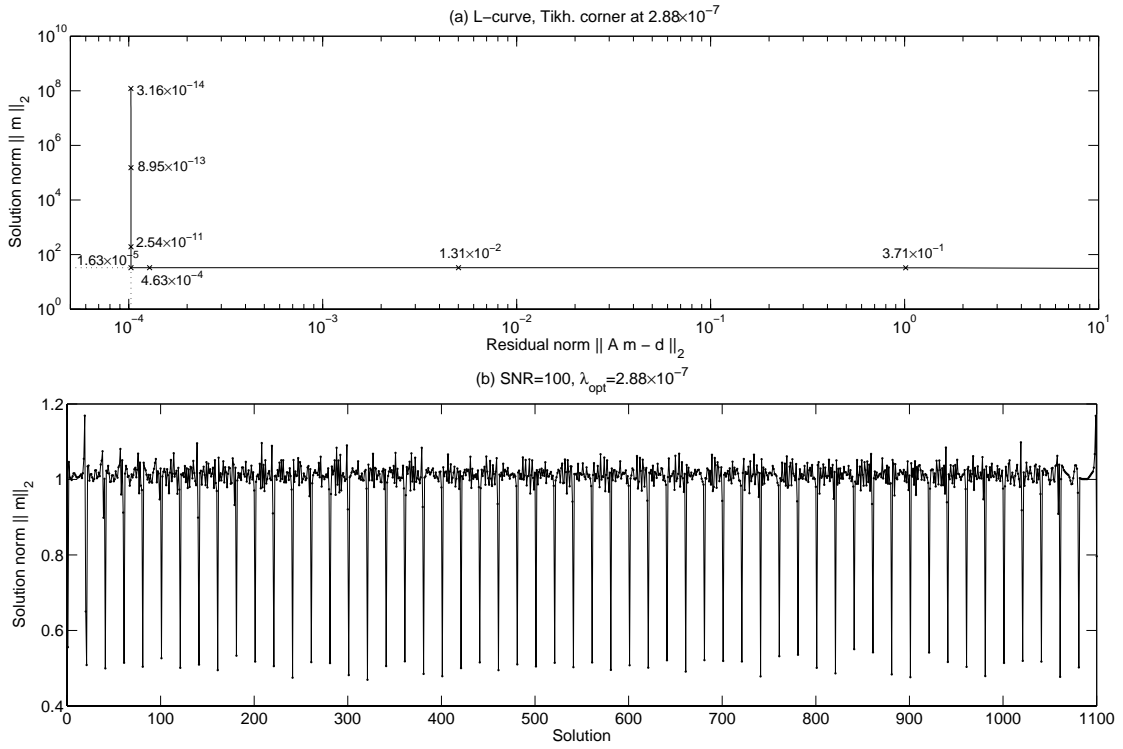


图 4-2 高信噪比(SNR=100)理论模型解算结果
(a:最佳正则化参数, b:理论模型的正则化解)

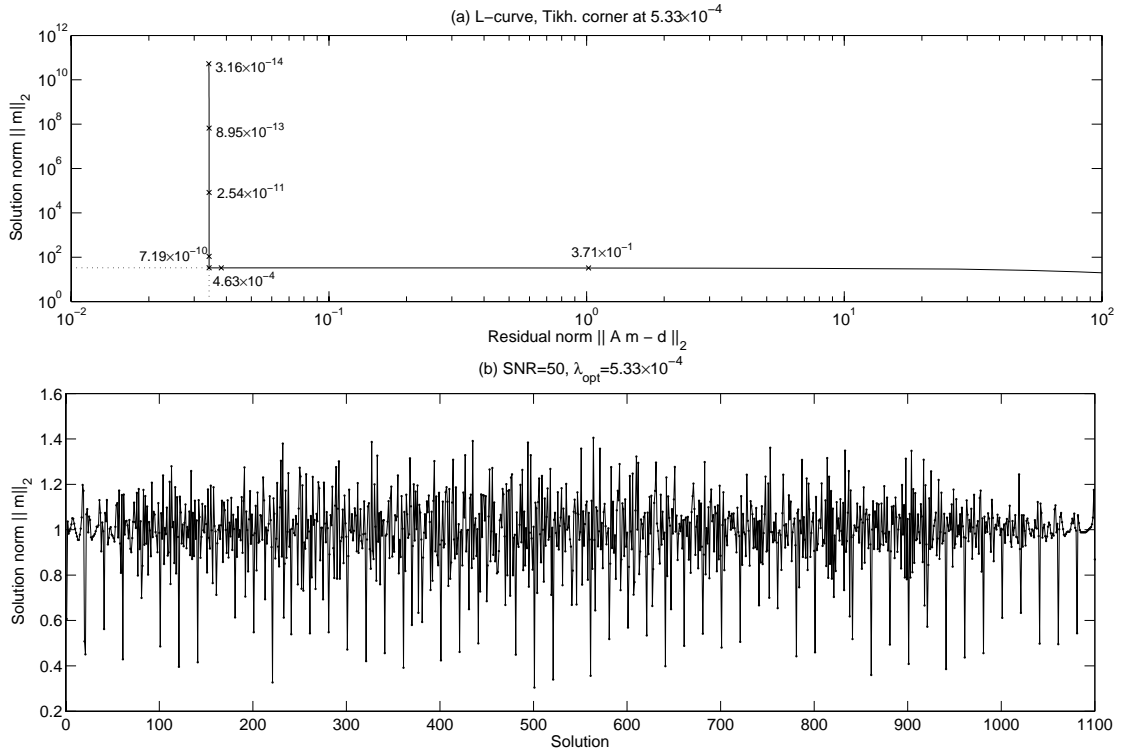


图 4-3 低信噪比(SNR=50)理论模型解算结果
(a:最佳正则化参数, b:理论模型的正则化解)

4.4 结果分析

(1) 同一成像方程、同一组观测数据, 采用不同的方法求解有可能获得不同的结果. 图 4-5、图 4-65、图 4-7 为对双频电磁波电导率层析成像采用不同的算法解算病态矩阵成像方程时获得的结果. 通过用本文方法及 LSQR、SP-LSQR、Tikhonov 正则化等方法反演得到的电导率图像与实际钻探资料和地质规律比较, LSQR 方法解算失效, SP-LSQR 在高噪反演、Tikhonov 正则化直接方法在低噪反演能得到较符合实际的解算结果, 而本文方法对高噪和低噪反演都能得到较满意的结果, 所得电导率成像边界清晰、成像图精细可靠.

通过对上述各种方法的分析认为, LSQR、SP-LSQR 反演结果不好的原因是缺少正则化项, 而这里的 Tikhonov 正则化方法是采用的拟牛顿法进行的优化计算, 在这里也较难获得稳定的反演结果.

(2) 传统的反演方法, 极少考虑反演参数的非负性, 本算法考虑了电导率参数的非负性, 更适合解决此类带非负约束的实际反演问题. 通过验证, 发现该正则化算法, 迭代稳定、抗噪声能力强, 特别适用于高噪数据反演.

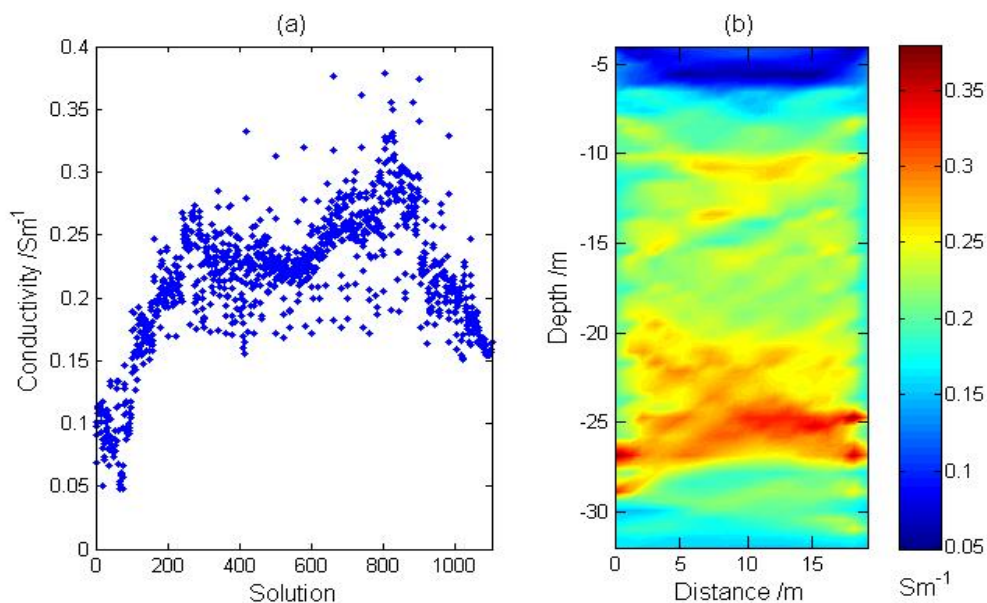


图 4-4 双频电磁波电导率成像 A-TR 反演结果（左）与重建的电导率图像(SNR=50)

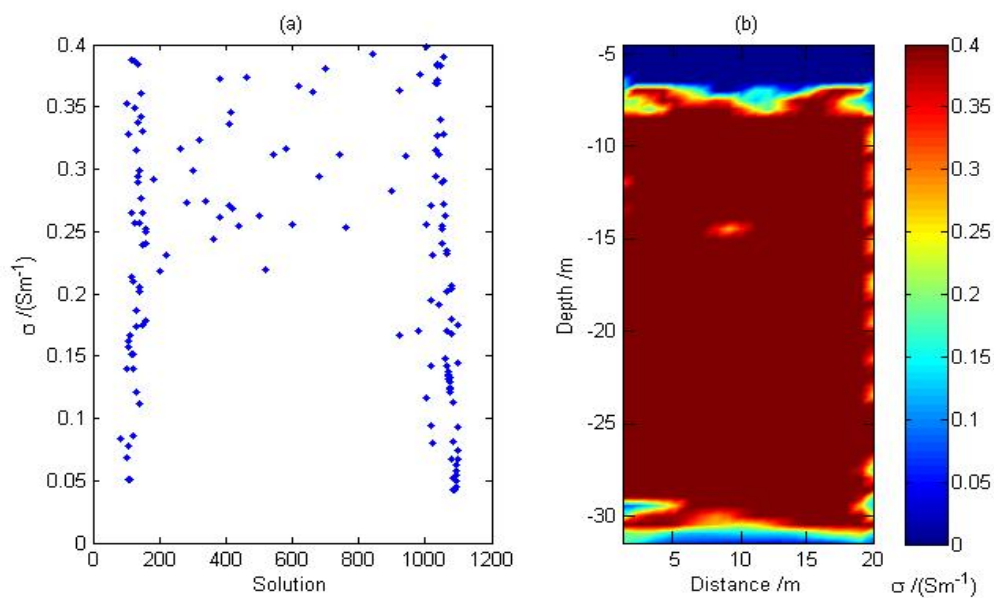


图 4-5 双频电磁波电导率成像 LSQR 反演结果（左）与重建的电导率图像(SNR=100)

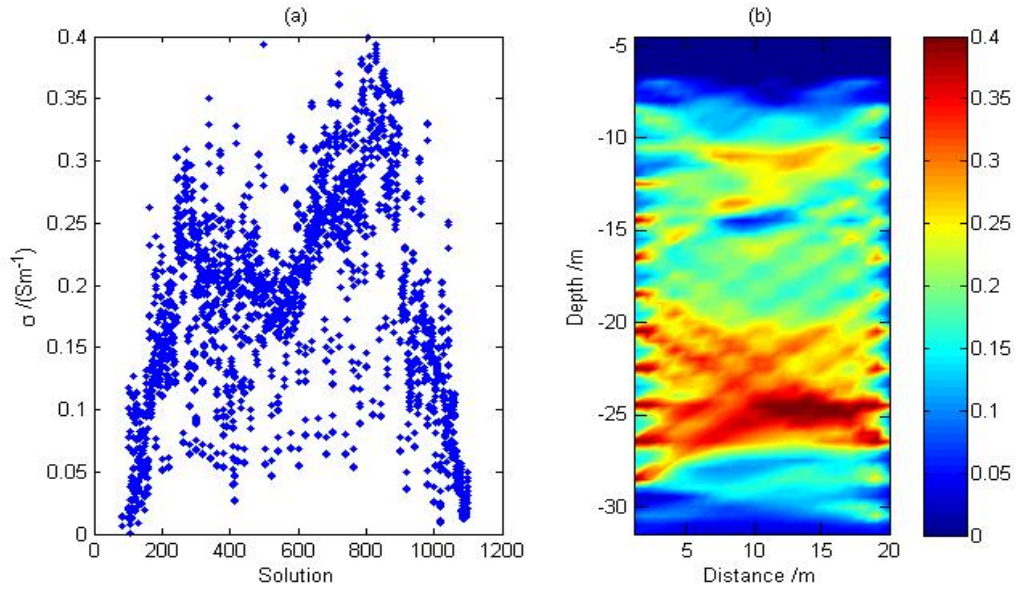


图 4-6 双频电磁波电导率成像 SP-LSQR 反演结果（左）与重建的电导率图像(SNR=100)

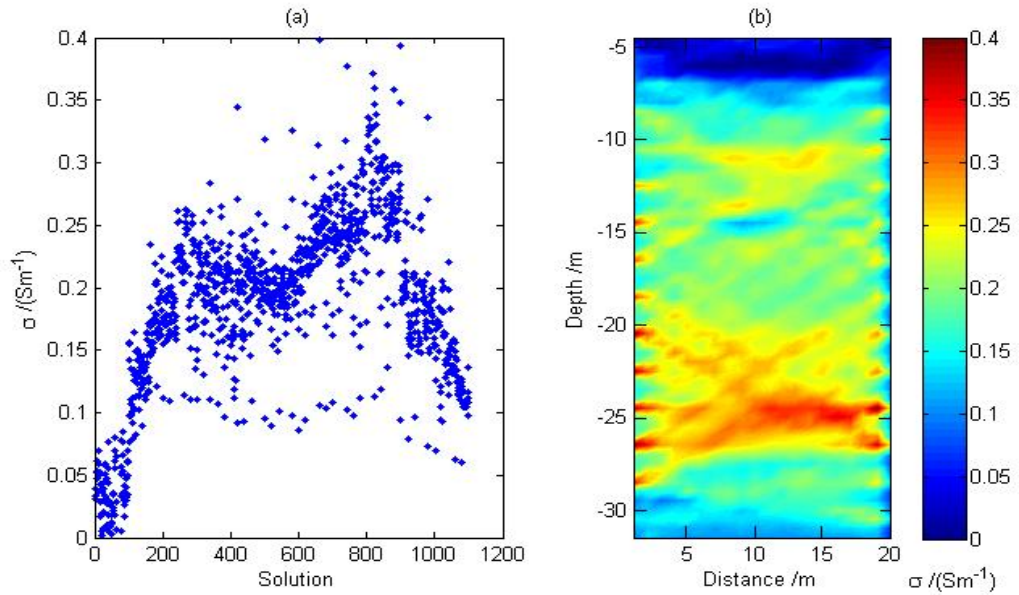


图 4-7 双频电磁波电导率成像 Tikhonov 反演结果（左）与重建的电导率图像(SNR=50)

第 5 章 基于差分进化算法的病态矩阵方程解算

在地球物理学中, 绝大多数的观测数据与模型参数之间都不满足线性关系, 作为这个领域的前沿, 自 20 世纪 90 年代以来非线性地球物理反演已经取得了很多可喜的进展^[94]. 在非线形反演方法中有一大类为线性化迭代方法, 这类方法多是从数学优化算法中引入的, 比较古老. 它的主要思想是将非线性问题线性化, 然后利用线性化迭代方法进行迭代计算. 其优点是迭代速度快, 不足之处在于易收敛于局部最优解、迭代结果依赖于初始点的选取和求解的模型参数等. 线性化迭代方法很多, 如最速下降法(the Steepest Descent method)、牛顿法(Newton method)、共轭梯度法(Conjugate Gradients, 简称 CG)、非线性最小二乘法(LSQR)等. 另一类非线性反演方法则是模拟自然界某种物理过程或现象的“启发式”反演法. 例如模拟“赌博”的蒙特卡洛法(Monte Carlo method)^[95,96]、模拟生物从低级到高级进化过程的遗传算法(Genetic Algorithm)^[97-99]、冷却结晶过程的模拟退火法(Simulated Annealing)、差分进化算法(Differential Evolution Algorithm, 简称 DE)、量子退火法(Quantum theory)^[100-102]和模拟原子跃迁过程的反演方法^[103]等. 这类非线性方法不依赖于初值和函数梯度值, 在收敛速度和避免陷入局部极值等方面具有一定的优势. 虽然非线性反演目前还处于蓬勃发展的阶段, 没有形成像线性反演那样完整的理论体系. 然而, 对于大多数为非线性、多极值的地球物理反演问题, 用非线性方法来解决必然是未来发展的趋势.

本章重点研究将 DE 算法应用于地球物理反演计算. 首先进行了将 Tikhonov 正则化方法与差分进化算法结合的电导率反演实验, 然后在如何加速差分进化算法的收敛速度方面展开研究, 提出了基于种群熵的自适应差分进化算法(简称 ARDE)以及粒子群差分进化混合算法(简称 PSODE), 将其应用于实际数据反演, 取得了较好的反演结果. 研究表明, 在不影响反演效果的前提下, 这两种方法均不同程度地提高了收敛速度, 降低了时间成本. 最后, 本章还进一步讨论了基于 LSQR 的混合差分进化算法, 该算法避开了 Tikhonov 和 TSVD 等算法正则化参数的选取困难. 相关成果见【104, 105】.

5.1 差分进化算法

差分进化(DE)算法是由 Rainer Store 和 Kenneth Price^[106]于 1996 年共同提出的一种基于实数编码的具有择优思想的全局优化进化算法, 其以遗传算法为基础, 在算法中引入扰动矢量, 利用个体间的距离和方向等信息进行搜索, 克服遗传算法容易“早熟”的缺陷, 是一种比较有发展前途的算法. 在 1996 年举行的

第一届国际 IEEE 进化优化竞赛上, 对提出的各种方法进行了现场验证, DE 被证明是最快的进化算法. 与传统的进化算法如遗传(GA)算法等相比较, 除了具有原理简单, 收敛速度快, 控制变量少, 无需设置初值和计算导数, 更易于理解和编程实现等特点之外, 其主要区别在于^[107]:

(1) 传统的 GA 算法采取二进制编码, 而 DE 算法采取实数编码;

(2) GA 通过两个父代个体交叉变异产生两个子个体, 而 DE 通过三个父代个体交叉变异产生一个子个体;

(3) GA 产生的子个体通过某种概率选择直接取代父个体, 而 DE 中新产生的子个体, 只有当它比相应的父代个体优良时才替换父个体.

差分进化(DE)算法由于其简单易用、具有可靠的鲁棒性而在短期内得到了很大的发展, 已经在人工神经网络、机器人、化工、图像处理等领域均得到了应用, 并迅速地被引入到众多科研和商业领域.

在反演方面, DE 算法也得到越来越多的应用. 潘克家等^[108]将 DE 算法应用于核磁共振 T_2 谱多指数反演, 韩伟等^[109]利用 DE 算法对油田地层参数进行反演, Harry Eckel 等^[110]将 DE 算法应用于 EIT 头部成像反演.

下面介绍差分进化(DE)算法的基本思想^[111].

DE 由 NP (种群规模) 个 D (决策变量个数) 维参数矢量 $x_{ij} (i=1, 2, \dots, NP; j=1, 2, \dots, D)$ 在搜索空间进行并行直接搜索. DE 算法包括变异 (Mutation)、交叉 (Crossover) 以及选择 (Selection) 三种基本操作, 最小化优化函数 f 的算法流程如图 5-1 所示.

为求非线性函数 $f(x_1, x_2, \dots, x_D)$ 的最小值, 进化过程中的第 G 代利用 NP 个 D 维参数向量构成种群 $\{x_{i,G}, i=1, 2, \dots, NP\}$, 种群大小 NP 在进化过程中保持不变.

(a) 初始化种群

在 D 维空间, 随机产生满足约束条件的 NP 个向量作为初始种群:

$$x_{ij}^0 = rand_i \cdot (b_{iu} - b_{il}) + b_{il}, \quad i \in \{1, \dots, D\}, j \in \{1, \dots, NP\}, \quad (5-1)$$

其中, $rand_i$ 为区间 $[0, 1]$ 上的随机数, b_{iu} 和 b_{il} 分别为第 i 个变量 x_i 的上下界, D 为参数的数量, NP 为设置的种群数量参数.

(b) 变异

对每一个目标向量 x_i^G 在 $\{1, 2, \dots, NP\}$ 范围内随机选择 3 个互异整数 r_1, r_2, r_3 且均与 i 不相等, 得到变异向量

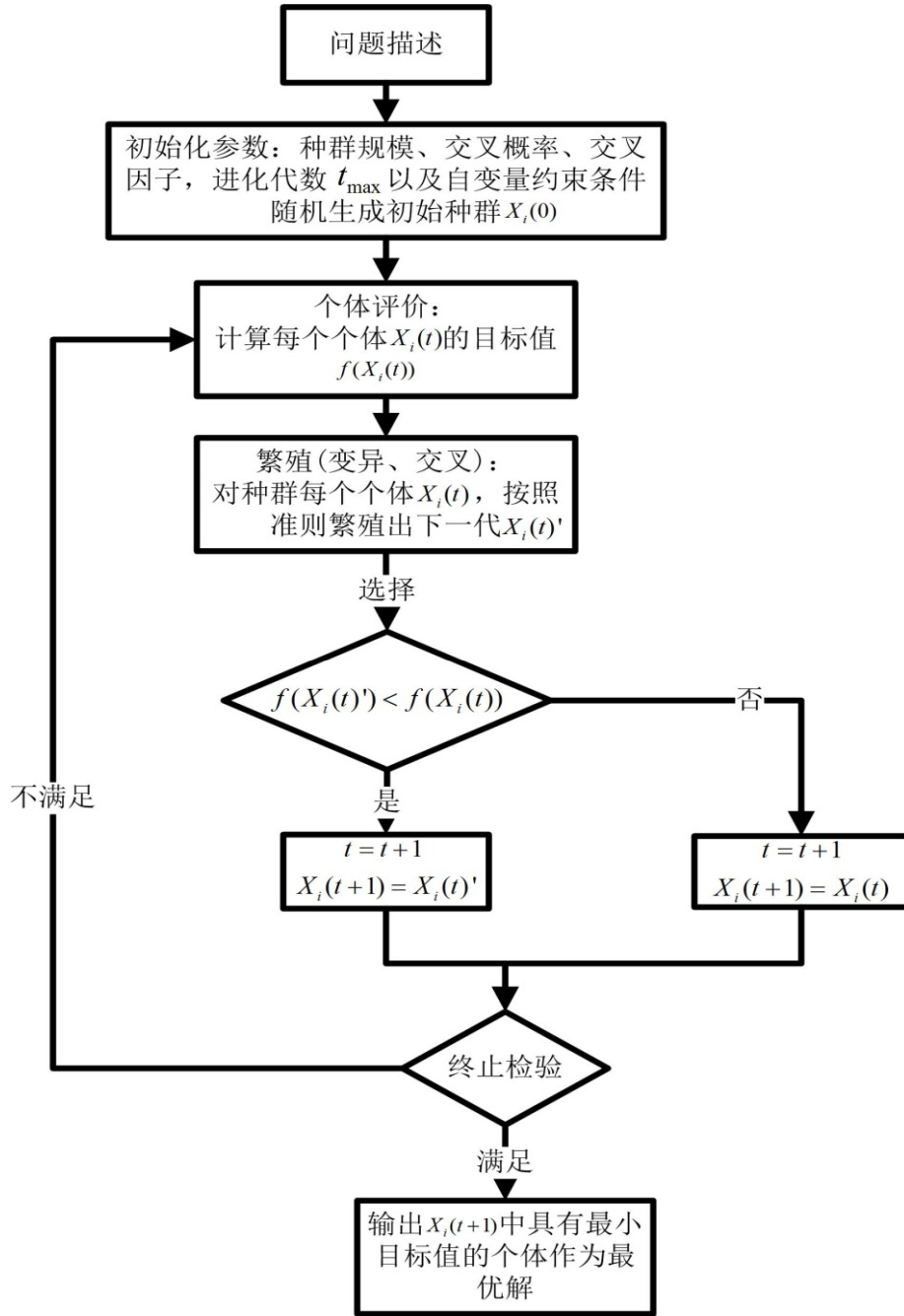


图 5-1 差分进化算法流程图

$$V_i^{t+1} = x_{r_3}^t + F \cdot (x_{r_1}^t - x_{r_2}^t), \quad (5-2)$$

其中, t 为进化代数, 变异因子 F 是介于 $[0,1]$ 之间的常数, 控制偏差向量的放大幅度.

(c) 交叉

交叉的目的是为了增加种群的多样性, 通过如下方式得到试验向量:

$$u_{ij}^{t+1} = \begin{cases} v_{ij}^{t+1}, & \text{rand}(j) \leq CR \text{ or } j = \text{randn}(i), \\ x_{ij}^t, & \text{rand}(j) > CR \text{ and } j \neq \text{randn}(i). \end{cases} \quad (5-3)$$

式中, $\text{rand}(j) \in [0,1]$ 为均匀分布的随机数; CR 为交叉概率常数, $CR \in [0,1]$, 本文算例中取作 1. 而 $\text{rand}(i) \in \{1, 2, \dots, D\}$ 保证该种群至少有一个基因要进行交叉变异.

在进行变异、交叉之后得到了试验向量, 通过比较试验向量和上一代种群的适应值, 选择得到适应值较优的参数作为下一代种群. 以最小化优化为例, 选择操作按照以下准则进行:

$$x_i^{t+1} = \begin{cases} u_i^{t+1}, & f(u_i^{t+1}) < f(x_i^t), \\ x_i^t, & f(u_i^{t+1}) \geq f(x_i^t). \end{cases} \quad (5-4)$$

算法将对种群参数进行边界值检验, 对于不在可行解区域中的参数, 将产生一个属于可行解区域的随机数将其替代.

差分进化算法原理简单, 受控参数少, 这使得参数的设置显得尤为重要. 合理的算法参数将有助于提高计算精度, 根据应用研究, 可得到以下一些经验性的结论^[112-115].

1) 种群数量 NP , 根据问题的规模, 一般选择在 $5D \sim 10D$ 之间. NP 过大将导致算法收敛速度较慢, 过小则可能造成 DE 算法过早收敛, 得不到目标函数的全局最小值;

2) 变异因子 F 是在变异操作中添加到被扰动向量上的差异值的比率. F 越大, 算法的全局搜索能力越强; 反之则局部搜索能力强;

3) 交叉因子 CR 表示子代个体从变异操作产生的向量实验个体中继承的比率. CR 较大, 收敛速度加快, 但易早熟.

5.2 基于差分进化算法的电导率成像反演

本节将利用基于全局寻优的差分进化算法对病态问题

$$A_{l \times n} m = d + e \quad (5-5)$$

进行求解. 利用 Tikhonov 正则化方法求解线性方程组(5-5), 将多次得到的近似解 x_k 组成 DE 算法的初始种群; 然后通过变异、交叉和选择等操作, 进一步优化所得到的正则化解.

首先做正则化处理:

$$\min(\|Am - d\|_2^2 + \lambda \|m\|_2^2) \quad (5-6)$$

其中正则化参数的确定采用 Hansen 创立的 L-曲线方法求解得到.

通过多次的理想模型反演,确定了以 $DE/rand/1/bin$ 为进化策略的最优差分进化模型,其差分进化参数为: $F = 0.6, CR = 0.7, NP = 6D$.

5.2.1 理论模型分析

在本节,先给定一组反演参数值 m_f , 假设 $m_f = (1, 1, \dots, 1)^T$, 得到式(5-5)的右边项 $d_f = Am_f$. 将 A 和 d_f 带入反演模型求解, 根据其反演结果 m 对 m_f 的逼近程度, 对反演模型的有效性进行验证.

对理想模型得到的 d_f 加上一定大小的高斯白噪音得到 \tilde{d}_f , 选取适合的正则化参数, 然后由数据对 (t, \tilde{d}_f) , 利用上述算法进行理想模型反演.

这里信噪比 SNR 仍然按如下定义:

$$SNR = 20 \log_{10} \frac{\|d\|_2}{\|d - \tilde{d}\|_2} \quad (5-7)$$

具体的方法是选择当 $SNR = 50$ 、 $SNR = 100$ 以及 $SNR = \infty$ 三种情况分别讨论. 首先通过利用 Tikhonov 正则化方法求解线性方程组(5-5), 每确定一个正则化参数, 就得到相应的一组初始解. 将多次得到的近似解 x_k 组成 DE 算法的初始种群; 再将这些近似解通过变异、交叉和选择等操作进行迭代求最优. 通过几种情况的比较发现, 差分进化算法具有较好的抗噪能力, 随着噪声的增加, 虽然相对误差也增加了, 但变化不大. $SNR = \infty$ 时, 相对误差为 0.0449, 而 $SNR = 100$ 和 $SNR = 50$ 时, 相对误差分别增加为 0.0528 和 0.0583.

通常, 正则化参数对于平衡观测值的扰动误差和引入正则化项带来的误差起着很重要的作用, 选取合适的正则化参数对于解的稳定性和有效性起着重大的影响^[116].

下面, 为了讨论 DE 算法对正则化参数的敏感性, 采用如下方法进行测试. 由于最优正则化参数的确定是一个非常困难的问题, 常常不容易找到最优正则化参数, 而这又影响着整个问题的精确解. 所以这里进行一个简单的测试. 选择正则化参数的不同数量级, 在不同数量级下选择不同的正则化参数, 确定了一系列的近似解, 用这些解再应用差分变异、交叉和选择等操作进行迭代求最优. 通过反演结果得出结论: 不同数量级的正则化参数下, 反演误差区别较大, DE 算法对正则化参数的敏感性较高.

通过上述讨论, 可以看出, 为了提高反演的精确性, 本文所设计的算法在解决此问题时是有效的.

5.2.2 双频电导率成像反演实验

经由上一节的讨论，得到 DE 算法抗噪能力强、对正则化参数敏感性高的认识，适合于应用到本例的实际应用当中。由 L-曲线法求得最优正则化参数为 8.0506，根据已经确定的 DE 算法模型，得到图 5-4 所示的反演结果。从反演结果看，DE 算法作为一种随机优化算法，能通过交叉、变异和选择等操作得到可信度较高的结果。

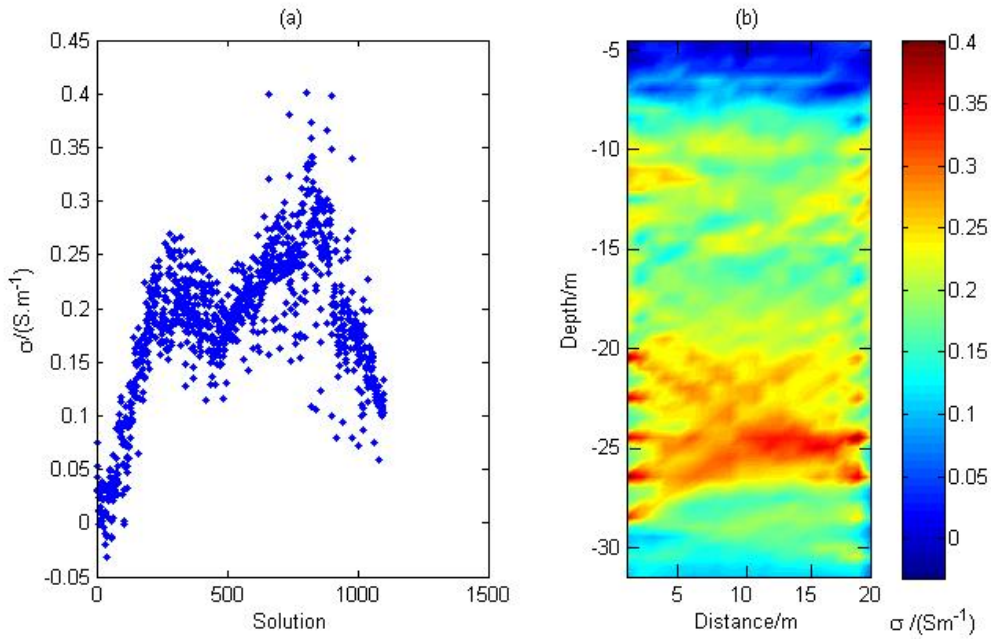


图 5-4 双频电磁波电导率成像 DE 反演结果（左）与重建的电导率图像

5.3 基于种群熵估计的自适应差分进化算法的解算方法

由于实际问题的反演参数常常较多，利用 DE 算法迭代求解的时间成本高，迭代次数受到较大的限制。而在较少迭代次数的条件下，很难获得接近于全局收敛的结果。因此，提高 DE 算法的收敛速度相当必要，这也是启发式算法一个很重要的研究方面。只有这样，在诸如本例的大型反演问题当中，才能以较小的时间成本取得满意的计算结果，使 DE 算法更好地满足实际工程的需求。

国内外学者先后提出了自适应 DE 算法以及混合 DE 算法等以改进 DE 算法的收敛性问题。为了提高差分进化的收敛速度，Josef^[117]提出了自适应差分进化算法(Adaptive Range Differential Evolution, 简称 ARDE)，并对其进行了对比研究。下面首先介绍一下自适应差分进化算法。并对 DE 算法进行改进，将其应用

于本例的实际反演当中。

5.3.1 自适应差分进化(ARDE)算法

如果单纯地增大交叉因子 CR 和变异因子 F ，一方面可以加快 DE 算法的收敛速度，但另一方面，却容易陷入局部最优及造成算法早熟。为了克服这个问题，提出了时变变异因子 F 及交叉因子 CR 的思想^[118,119]，也就是使变异因子 F 随迭代次数减小，交叉因子 CR 随迭代次数变大，使其在初期全局搜索能力强，而在后期则提高局部搜索能力。本节将时变变异因子 F 加入到 ARDE 算法中进行计算。

定义 5-1^[120,121] 设种群规模为 M ， f_{\min} 和 f_{\max} 是搜索过程中发现的最大和最小适应值，将目标函数的解空间划分为 M 个不同区域 A_i , $i=1,2,\dots,M$ ：则

$$A_i = [f_{\min} + (i-1) \frac{f_{\max} - f_{\min}}{M}, f_{\min} + i \frac{f_{\max} - f_{\min}}{M}] . \quad (5-8)$$

设种群中有 M_i 个染色体的适应值属于 A_i ，定义染色体出现在第 i 个区域的概率为

$$p_i = \frac{N_i}{N} , \quad (5-8)$$

定义第 t 代种群的种群熵为

$$S_t = - \sum_{i=1}^M p_i \ln p_i , \quad (5-10)$$

定义改变种群范围的控制参数 S_r 为

$$S_r = 0.98 \exp \frac{-t}{t_{\max}} , \quad (5-11)$$

其中， t 为种群的迭代代数， t_{\max} 为设定的最大迭代代数。设 S 为初始化种群的熵，设变量的取值区间 $x_j \in [r_j^{\min}, r_j^{\max}]$, $j=1,2,\dots,N$ ， N 为变量的数量， $GN_{t,j}^{\min}, GN_{t,j}^{\max}$ 分别为第 t 代种群中适应值前 60% 个体的第 j 个变量的最小、最大值。当 $S_t/S < S_r$ 时，改变种群搜索范围：

$$\begin{aligned} r_j^{\min} &= w GN_{t,j}^{\min} + (1-w) r_{tj}^{\min} \\ r_j^{\max} &= w GN_{t,j}^{\max} + (1-w) r_{tj}^{\max} \end{aligned} \quad (5-12)$$

其中 w 为松弛因子，介于 $[0,1]$ 之间。根据 S_r 的定义，可以看出 S_r 在搜索初期较大，种群的搜索范围会以较快的速度减小，但是随着迭代的进行， S_r 逐渐变小，种群自适应减小搜索区间的频率减小。

为了使差分进化(DE)算法前期具有较好收敛速度，后期能够稳定地收敛，在进化的过程中，将交叉因子 F 定义为随迭代代数线性变化，即

$F = F_{\max} - \frac{(F_{\max} - F_{\min})}{gen}i$, 其中 gen 为最大迭代次数, i 为迭代的次数, F_{\max}, F_{\min} 分别为设置的最大、最小交叉因子。

5.3.2 基于 ARDE 算法的电导率反演及结果分析

为了优化差分进化算法的搜索能力, 随算法迭代的进行, 自适应减小变异因子 F , 设定 $F_{\max} = 0.9$, $F_{\min} = 0.3$, 算法中其他参数和设置不变。

ARDE 在前期由于寻优区间收缩频率比较大, 收敛速度比较快, 后期由于 S_r 较小, 收敛趋于平稳。图 5-5 为自适应差分进化算法(ARDE)和标准 DE 算法的收敛比较图, 从图中可以看出, ARDE 在 200 代时就已经达到了 DE 算法在 500 代的收敛水平, 表明 ARDE 算法具有实际的应用价值。

基于种群熵差分进化算法经过 250 次迭代得到的实际反演结果如图 5-6。

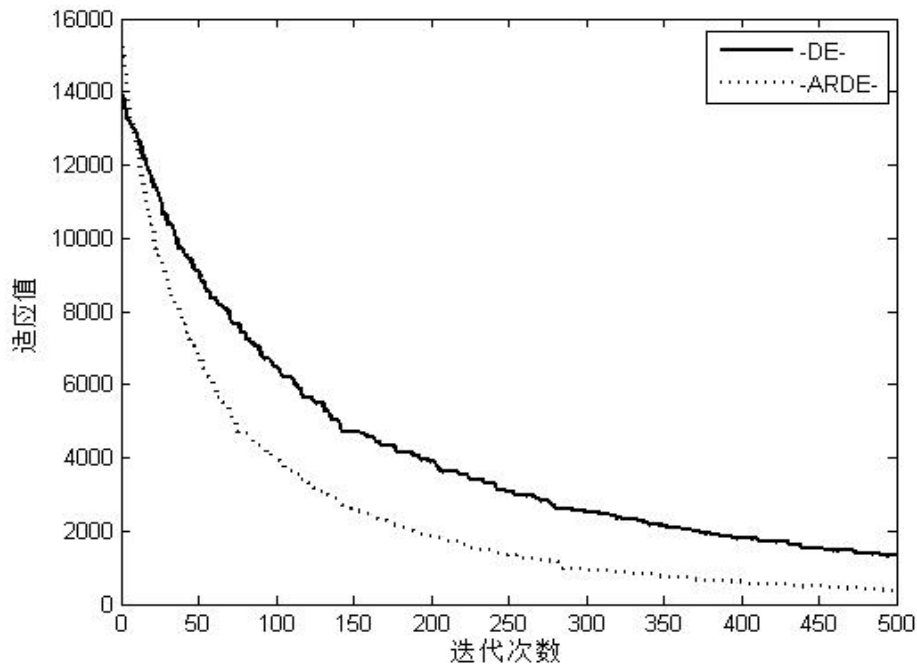


图 5-5 DE 算法和 ARDE 算法的收敛比较图

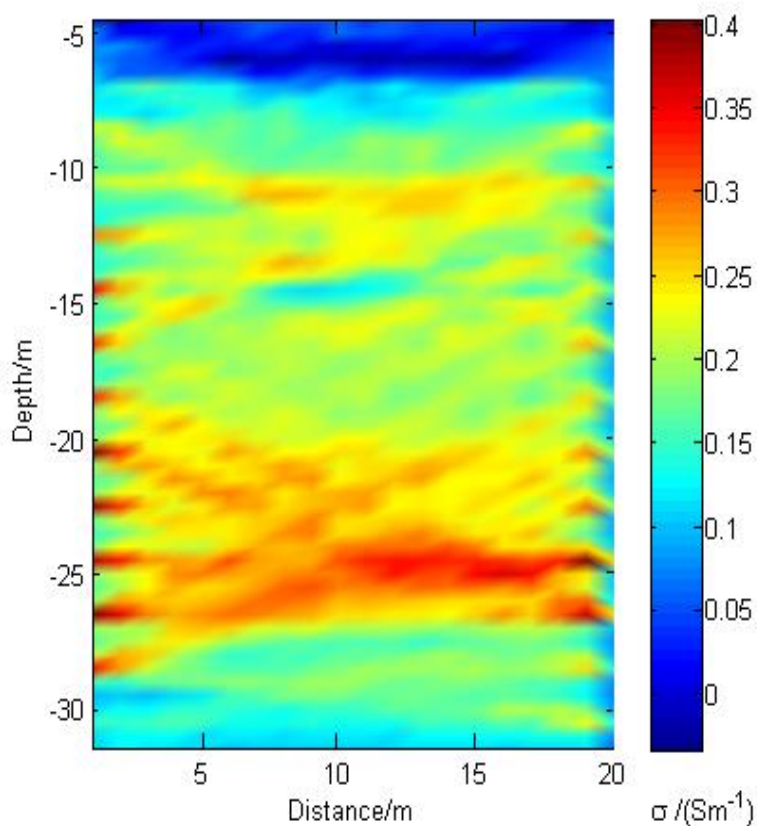


图 5-6 双频电磁波电导率成像 ARDE 算法重建的电导率图像

5.4 基于粒子群差分进化算法的解算方法

在计算智能领域有两种基于群集智能的算法：一种是蚂蚁算法(Ant Colony Optimization, 简称 ACO)，另一种则是粒子群算法(Particle Swarm Optimization, 简称 PSO)。PSO 算法是在 1995 年由美国社会心理学家 Kennedy 和电器工程师 Eberhart^[122]共同提出的，它借鉴了鸟群或鱼群捕食过程的社会行为，是一种有别于遗传算法(Genetic Algorithm, GA)的并行进化计算技术。在 1997 年他们又提出了二进制 PSO 算法^[123]。Angeline^[124,125]于 1998 年借鉴进化计算中的选择概念，将其引入 PSO 算法中，而 Lovbjerg 等^[126]进一步将进化计算机制应用于 PSO 算法。

5.4.1 粒子群(PSO)算法

下面介绍 PSO 算法思想^[127]：在 PSO 算法中，每个优化问题的解看作是一个个粒子，所有的粒子都有一个由被优化的函数决定的适应值，每个粒子还有一个速度决定其进化方向和距离，粒子群中的粒子追随当前的最优粒子在解空间中搜

索。每一次迭代中，粒子通过跟踪两个“极值”来更新自己。一个是粒子本身所找到的最优解叫做个体极值 $pbest$ ，另一个极值是整个种群目前找到的最优解，这个极值是全局极值 $gbest$ 。找到这两个极值，每个粒子根据如下公式来更新自己的速度和新位置：

$$v_{k+1} = c_0 v_k + c_1 (pbest_k - x_k) + c_2 (gbest_k - x_k) \quad (5-13)$$

$$x_{k+1} = x_k + v_{k+1} \quad (5-14)$$

其中， v_k 是速度向量， x_k 是当前粒子的位置。 c_0, c_1, c_2 是种群的认知系数： c_0 为惯性权重，一般介于 0 和 1 之间， c_1 和 c_2 是加速常数，介于 0 至 2 之间。一般情况下，会对更新速度 v 限定在一定的范围 v_{\max} 内，即当 $v_k > v_{\max}$ 时， $v_k = v_{\max}$ ，当 $v_k < -v_{\max}$ 时， $v_k = -v_{\max}$ 。其算法流程描述如下^[128]：

(1) 初始化粒子群：预设迭代次数 $Time$ ，设定粒子数量 n ，惯性权重 c_0 和加速常数 c_1, c_2 ，设置最大速度 v_{\max} ；

(2) 随机产生 n 个粒子的初始位置 X_0 及其初始速度 V_0 ，计算各个粒子的适应值，根据适应值找到个体极值 $pbest$ 以及全局极值 $gbest$ ：

$$\begin{aligned} pbest_i &= X_0(i), \\ pval_i &= f(pbest_i), \end{aligned} \quad (5-15)$$

其中， $pval_i$ 为第 i 个粒子的适应值， f 为计算适应度的目标函数，假设 $pval_k = \min(pval_i)$ ，则 $gbest = pbest_k$ ， k 为最佳适应值粒子的索引号；

(3) 按照(5-13)更新自己的速度，并将其限制在 v_{\max} 内，根据(5-14)更新当前位置，得到新的粒子群 X ；

(4) 对于每一个粒子，计算其适应值，当其适应值小于 $pbest$ 计算所得的适应值时，更新 $pbest$ ，即：

$$pbest_i = X(i), \quad \text{当 } f(X(i)) < f(pbest_i); \quad (5-16)$$

反之，则不予更新。对于新粒子群，按照(2)中的公式更新 $gbest$ ，返回(3)，直至达到迭代精度或者达到迭代次数，则退出循环。

在算法中， X 为 $m \times n$ 的矩阵， m 表示粒子种群的数量， n 为目标函数参数的个数； $pbest$ 亦为 $m \times n$ 的矩阵，每一行记录了对应粒子截至迭代当次的最优解； $gbest$ 为 $1 \times n$ ，记录的是所有粒子截至迭代当次的最优解，显然有 $f(pbest_i) \leq f(gbest), i \in [1, m]$ 。

粒子群算法(PSO)在求解最优化问题的初期具有收敛速度较快，后期由于所有粒子均向最优粒子靠近，群体失去了多样性，粒子易陷入局部最优。

5.4.2 粒子群混合差分进化 (PSODE)算法

DE 算法是通过变异、交叉、选择产生适应性的后代，而 PSO 算法则是不断更新粒子的位置和运动方向以使其不断逼近最优解。为了加速差分进化算法(DE)的收敛速度，同时避免陷入局部最优的困境，将 DE 算法和粒子群算法(PSO)结合，使它们互相克服彼此的算法缺陷。文献【129.130】提出了基于 PSO 算法和 DE 算法的混合全局优化算法(PSODE)。下面，对 PSODE 算法进行描述^{[129]错误！未找到引用源。}。

(1) 初始化，首先确定种群规模的 $popsiz$ e、迭代的次数 t ，最大迭代次数 T_{\max} 、最大惯性权重 C_{\max} 、最小惯性权重 C_{\min} 、最大速度 V_{\max} 、控制因子 ϕ 、加速因子 C_1, C_2 ，缩放因子 F 以及变异概率 CR ；

(2) 将种群等分为两个种群 POP^{PSO} 和 POP^{DE} ，分别对两个种群进行初始化，设置迭代计数器 $t=0$ ；

(3) 利用(式 5-13)、(式 5-14)以及

$$C_0 = C_{\min} + (C_{\max} - C_{\min}) \exp \left[-\phi \left(\frac{t}{T_{\max}} \right)^2 \right] \quad (5-17)$$

对 POP^{PSO} 种群中所有个体进行位置、速度更新；

(4) 根据(式 5-2)、(式 5-3)、(式 5-4)对 POP^{DE} 种群中的个体执行变异、杂交、选择操作；

(5) 分别选出 POP^{PSO} 和 POP^{DE} 种群中的最佳个体 g_{best}^{PSO} 以及 g_{best}^{DE} ，比较其优劣，选择最佳个体作为 POP^{PSO} 和 POP^{DE} 下一代进化的依据；

(6) 对每一个个体，根据(5-16)判断是否出现停滞现象，若出现，则按照

$$X_i^{t+p+1} = X_{\min} + rand(X_{\max} - X_{\min}) \quad (5-18)$$

执行变异操作，其中 X_{\max}, X_{\min} 是定义的搜索区间；如果没有，返回执行步骤(3)至步骤(5)，直至达到最大迭代次数，则终止程序。

5.4.3 基于 PSODE 算法的电导率反演及结果分析

图 5-7 为 PSO 算法、DE 算法、PSODE 算法进行理想模型反演的收敛速度图，不难看出，PSO 算法迭代前期具有较快的收敛速度，随着进化种群逐渐失去多样性，粒子陷入局部最优。PSO 算法在迭代初期高效的收敛能力优化了 DE 算法的输入，使 DE 算法与 PSO 算法结合得到的 PSODE 算法，收敛速度得到了大幅提高。

图 5-8 为 PSODE 算法经 250 次迭代得到的反演成像效果图，显示 PSODE 算法具有较快的反演速度，同时也保持了较好的反演精度，反演结果的 l_2 范数误

差为 1.08%.

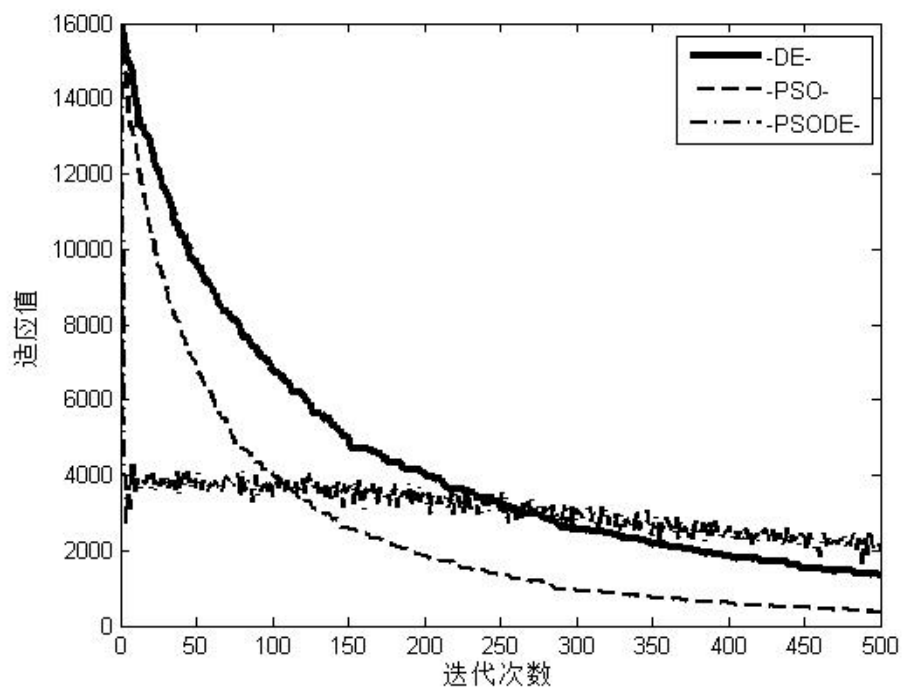


图 5-7 DE、PSO 、PSODE 算法的收敛比较图

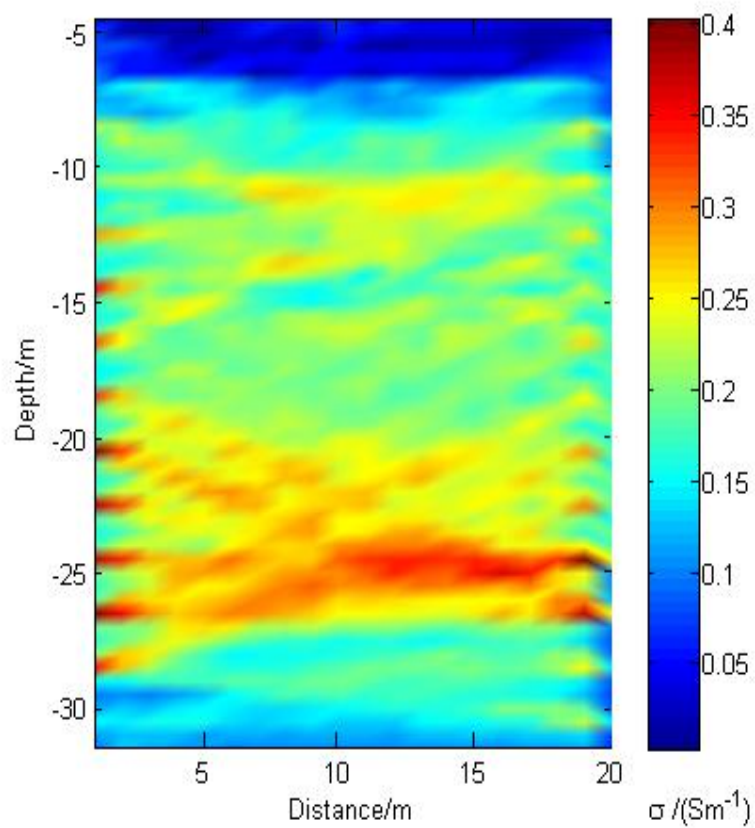


图 5-8 双频电磁波电导率成像 PSODE 算法重建的电导率图像

5.5 基于 LSQR 的混合差分进化算法的解算方法

在第 4 章中提出基于 Tikhonov 正则化的 Active-Set 算法,该算法虽然在保证了所求得电导率为非负的物理限制条件下,克服了大型病态方程的求解困难,但是却始终无法避开正则化参数的选取.而在许多实际反演计算中,正则化参数的有效确定一直是一个非常困扰的问题.

这一节所要介绍的基于 LSQR 的混合差分进化算法(简称 HDE 算法)是一种在迭代正则化的基础上采用 DE 进化策略的优化新算法^[104].该算法避开了 Tikhonov 和 TSVD 等算法正则化参数的选取困难,具有数值稳定性好、不依赖于初值、不易陷入局部极值和收敛速度快等优点.

具体实现步骤为:

(1) 利用 LSQR 方法求解线性方程组(5-5),将各步迭代得到的近似解 x_k 组成 DE 算法的初始种群;

(2) 通过变异、交叉和选择等操作,利用 DE 算法求解二次优化问题 $\min_m \|Am - d\|^2$,进一步优化所得到的正则化解.

图 5-9 为利用 HDE 方法求得其成像方程的离散解再插值后重建的电导率图像.本节仅以双频电磁波电导率成像为例,来验证了 HDE 算法求解大型病态线性问题的可行性.而对于该算法是否可以用来反演规模更大、病态性质更严重的其他地球物理线性问题,还需要进一步研究.

而对于该算法能够避开通常正则化方法最佳参数的选取困难这一特点,无疑将是此方法能够应用于一些特殊地球物理反演问题最大的优势.

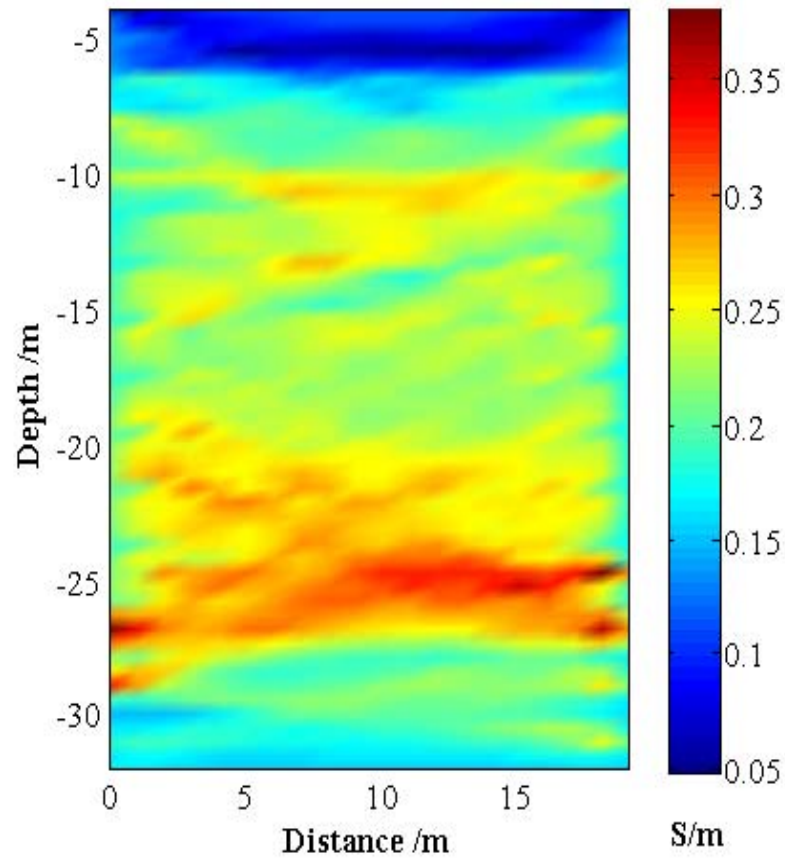


图 5-9 双频电磁波电导率成像 HDE 算法重建的电导率图像

第6章 双参数混合正则化方法的病态矩阵方程解算

本章通过引入带有二阶正则算子的正则化项,建立一种双参数混合正则化方法,并应用 L-曲线法、偏差原理和广义交叉校验准则的优化组合来确定最佳正则化参数,并以电导率反演成像为例,进一步研究正则化计算方法.通过对理论模型的数值模拟以及具体数据验证,表明了该方法的有效性.

6.1 双参数混合正则化解算方法

文献[131,132]在研究带限制的线性不适定问题时引入了双参数的正则化方法,文献[133]讨论了基于双参数正则化的非线性不适定问题,并研究证明了关于此问题解的存在性、稳定性和收敛性.

尽管数学家在理论上对双参数的正则化方法做了深入的讨论,但是地球物理学家却极少将此方法应用于实际的反演计算中.这主要是由于随着正则化项的增加,正则化参数的确定显得尤其复杂.下面我们将对此进行初步的探索与尝试.

6.1.1 双参数混合正则化方法原理

为求得方程

$$Ax = y + e \quad (6-1)$$

的一个更加稳定的近似解,并且减少多解性,我们基于标准形式的 Tikhonov 正则化方法,通过引入带有二阶正则算子的正则化项,设计了带双参数的正则化方法算法,即取下述极小化问题

$$\min_{x \in X} \left\{ \|Ax - y\|_2^2 + \alpha \|Lx - Lx_{reg0}\|_2^2 + \beta \|x\|_2^2 \right\} \quad (6-2)$$

的一个解作为方程(6-1)的近似解.第一个正则化项 $\|Lx - Lx_{reg0}\|_2^2$ 可以看作是引入了对参数 x 的约束,其中 x_{reg0} 是参数的先验值,可由截断奇异值分解法、共轭梯度法及 Tikhonov 正则化直接法分别求解后再进行联合优化所确定; L 为离散化后与正则算子相对应的正则矩阵,在此取为二阶正则矩阵

$$L = \begin{bmatrix} -1 & 2 & -1 & & \\ & -1 & 2 & -1 & \\ & & \ddots & \ddots & \ddots \\ & & & -1 & 2 & -1 \end{bmatrix}.$$

第二个正则化项取为 $\|x\|_2^2$. 这里 $\alpha > 0$, $\beta > 0$ 分别为正则化参数,它们的选取分别由第3章中曾经介绍的 L -曲线准则、Morozov 的偏差原理及广义交叉校验准

则分别进行计算，然后再进行优化组合，以此得到最佳正则化参数。

6.1.2 算法可行性分析

为了验证本文提出的双参数混合正则化方法的优越性，我们进行理论模型试验。以 Hilbert 矩阵作为系数阵的矩阵方程

$$\begin{bmatrix} 1 & 1/2 & 1/3 & \cdots & 1/n \\ 1/2 & 1/3 & 1/4 & \cdots & 1/(n+1) \\ 1/3 & 1/4 & 1/5 & \cdots & 1/(n+2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1/n & 1/(n+1) & 1/(n+2) & \cdots & 1/(2n-1) \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \cdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \cdots \\ y_n \end{bmatrix} \quad (6-3)$$

是不适定的。我们在(6-3)式中选择了 $n=101$ ，这时(6-3)式是一个高度不适定方程组，其条件数高达 2.0459×10^{19} 。在试验中，我们对右端项 y 加入噪声，假定该噪声具有可加性，则

$$y = y_{true} + \delta \times rand(size(y_{true})), \quad (6-4)$$

其中， δ 为噪声水平， $rand(size(y_{true}))$ 具有与 y_{true} 维数一致的 Gauss 噪声，噪声水平在试验中分别取为 0, 0.1 和 0.3。我们通过双参数混合正则化方法与第三章所述的截断奇异值分解法、共轭梯度法、Tikhonov 正则化直接解法分别进行反演计算比较，计算结果如图 6-1、图 6-2、图 6-3 和图 6-4 所示。通过比较发现，反演测试结果表明双参数混合正则化方法不仅具有较高的反演计算精度，而且对于数据的随机噪声是稳定的。

6.2 基于双参数混合正则化方法的电导率反演成像实验

6.2.1 理论模型分析

首先假设双频电导率反演问题中的电导率参数 $m_f = 1$ ，得到右端项 $d_f = Am_f$ ，然后在 d_f 上加入误差水平分别为 0, 0.1 和 0.3 的白噪声，得到 d_1 ， d_2 和 d_3 ，再利用双参数混合正则化算法反演得到电导率 m ，并和假定的电导率 m_f 进行比较，计算结果分别如图 6-5 所示。

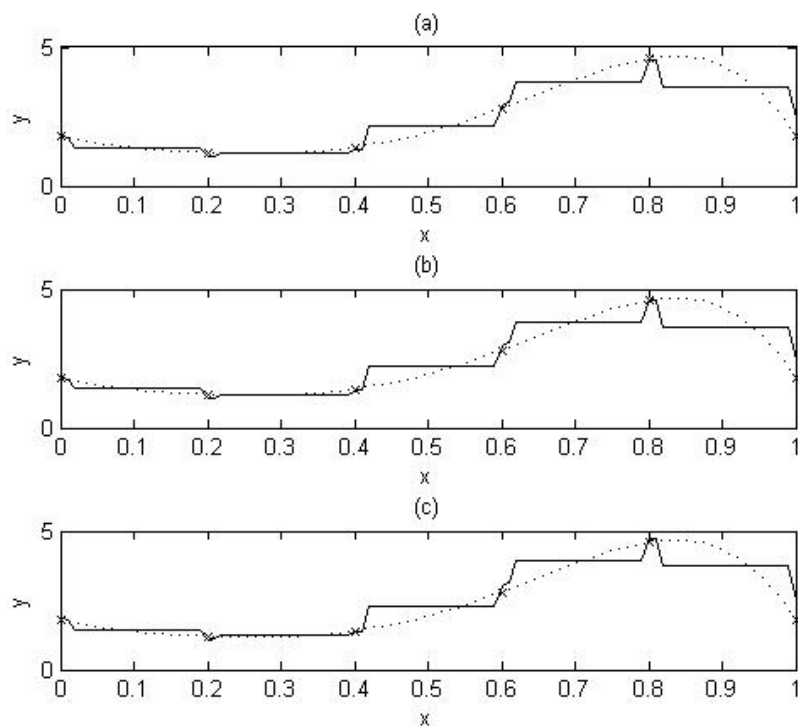


图 6-1 噪声水平为 0(a), 0.1(b)和 0.3(c)时的真实解 (虚线) 和截断奇异值分解法的解 (实线)

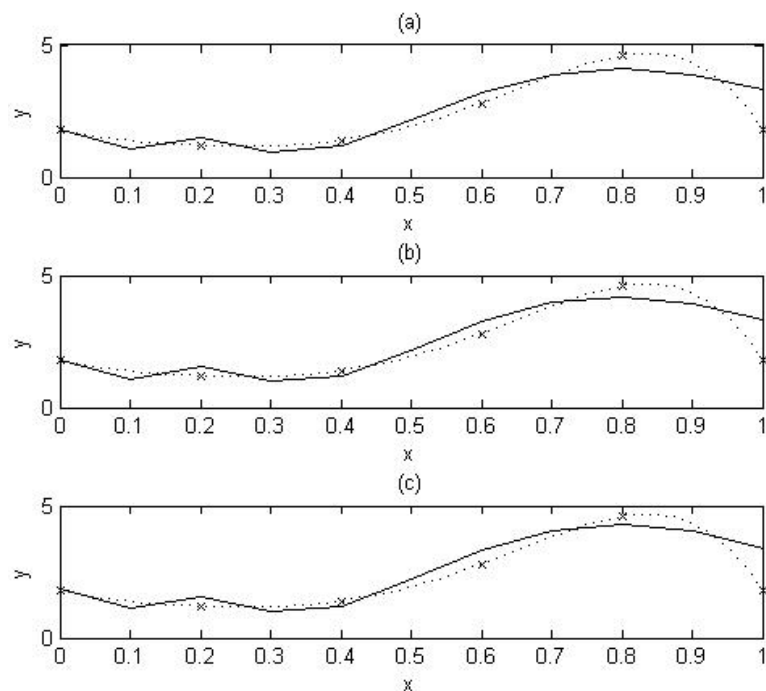


图 6-2 噪声水平为 0(a), 0.1(b)和 0.3(c)时的真实解 (虚线) 和共轭梯度法的解 (实线)

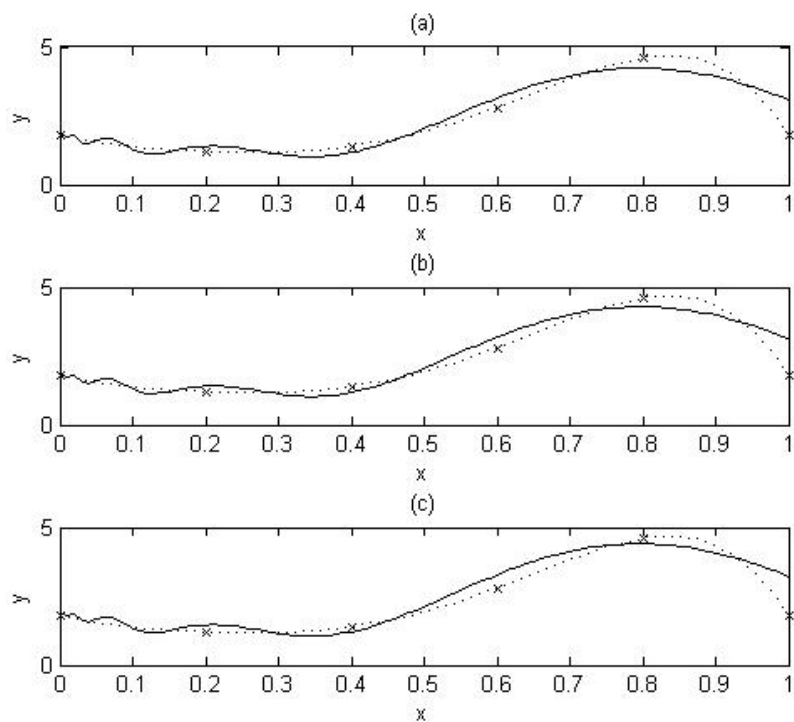


图 6-3 噪声水平为 0(a), 0.1(b)和 0.3(c)时的真实解（虚线）和标准 Tikhonov 正则化解（实线）

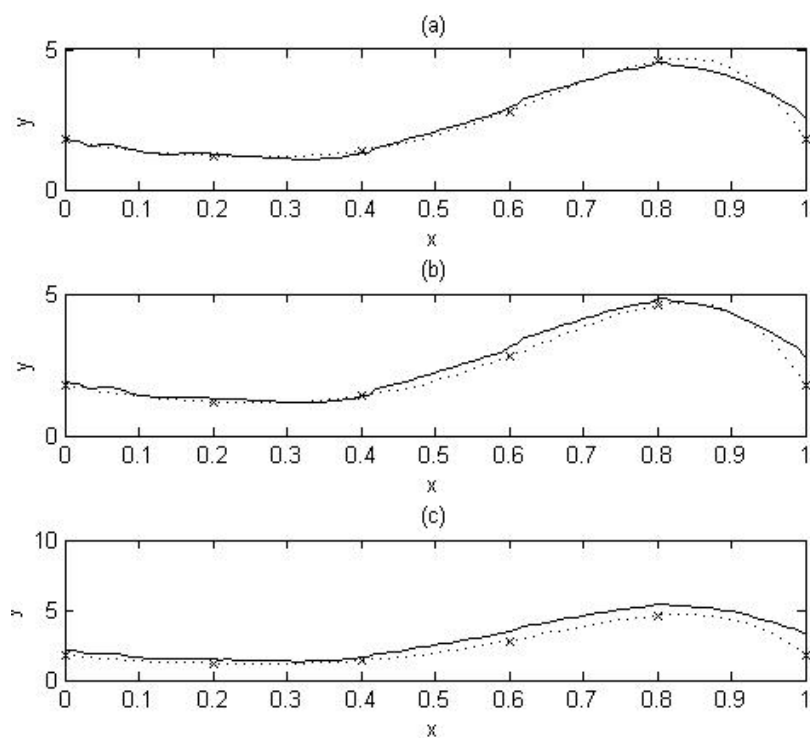


图 6-4 噪声水平为 0(a), 0.1(b)和 0.3(c)时的真实解（虚线）和双参数混合正则化解（实线）

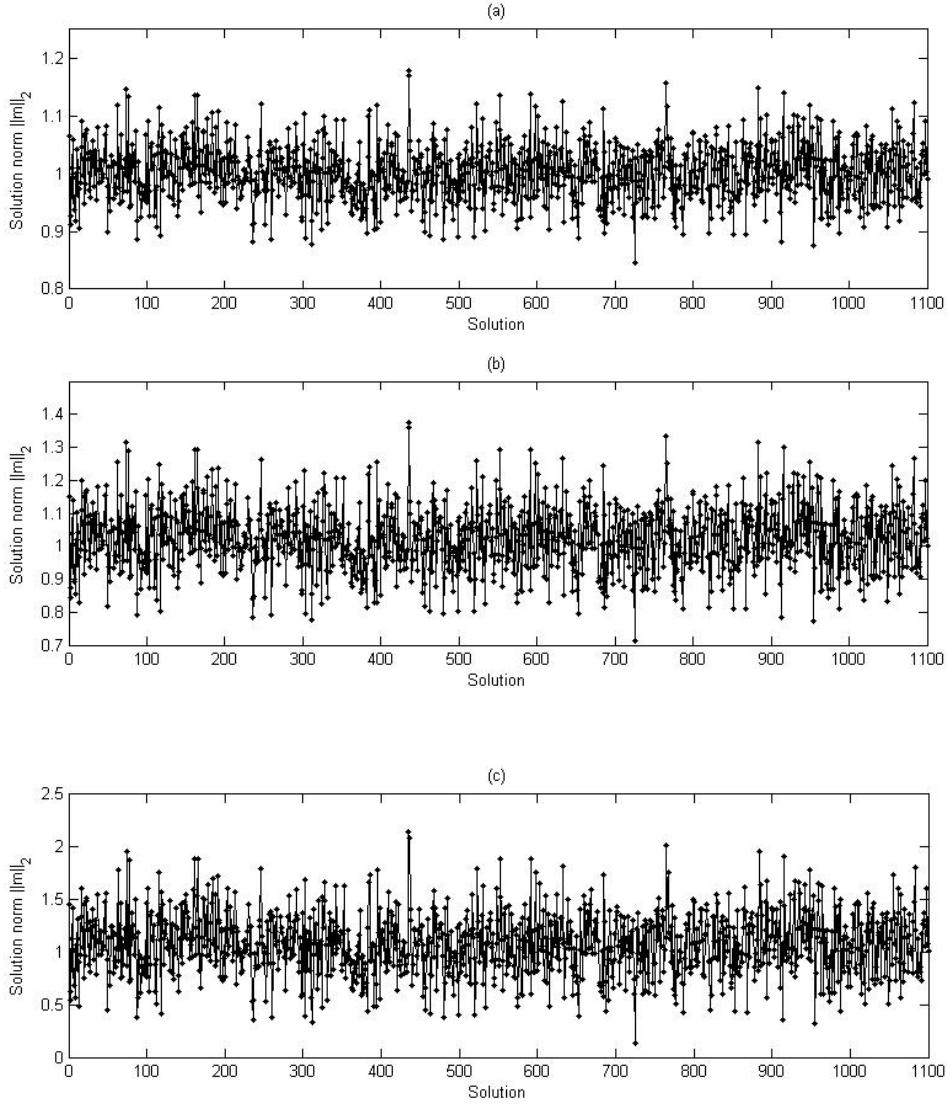


图 6-5 噪声水平为 0(a), 0.1(b)和 0.3(c)时的解

在第一个算例，误差水平为 0，引起的右端项误差 $\|e\|_2$ 为 0，最佳正则化参数分别为 3.22 和 3.84，正则化解 $m_{\lambda_{opt}}$ 在 2-范数下的相对误差为 4.97%；在第二个算例中，误差水平增大为 0.1，右端项误差 $\|e\|_2$ 增至 5.19×10^{-3} ，最佳正则化参数分别增至 5.52 和 6.18，所得的正则化解相对误差随之增大到 10.22%；在第三个算例中，误差水平增大为 0.3，右端项误差 $\|e\|_2$ 增至 1.56×10^{-2} ，最佳正则化参数分别增至 8.15 和 8.56，所得的正则化解相对误差随之增大到 15.65%。此三个算例说明了本文算法对这种问题求解的有效性，即使误差比较大（误差水平为 0.3），本文的正则化算法仍能得到比较好的反演结果。

6.2.2 双频电导率成像反演验证

通过截断奇异值分解法、共轭梯度法、标准 Tikhonov 正则化法和双参数混合正则化方法分别进行反演计算地层电导率。反演结果分别如图 6-6、图 6-7、图 6-8 和图 6-9 所示。上述几种正则化方法相比较，成像结果表明用双参数混合正则化方法得到了精确细致可靠的电导率成像图。成像图所反映的电性结构特征与钻探所揭露的剖面地质特征基本一致，连续性也较好，与钻探资料和地质规律吻合。

6.3 结果分析

(1) 本章提出的双参数混合正则化方法是一种地球物理反演计算的新的尝试。通过与截断奇异值分解法、共轭梯度法、标准 Tikhonov 正则化法相比较，理论模型数值反演结果表明，本文建立的正则化方法，对于数据的随机扰动有良好的稳定性，应用于电导率反演，成像细致可靠，也符合实际情况。

(2) 正则化方法是求解决不适定问题的有效方法，而通过选取不同的正则矩阵可以改变正则化项所包含的先验信息，可以改善问题的病态性，实例表明某些情况下，引入带有二阶正则矩阵的正则化项可以保证反演成像的精确度。

(3) 在实施正则化方法过程中，正则化参数的选取是至关重要的，尤其是本文中的正则化项有两项的情况下。本章是通过 L 曲线准则、Morozov 的偏差原理及广义交叉校验准则的联合计算得到最佳正则化参数，增加了计算工作量，在确定正则化参数方面还需要研究更加有效的方法。

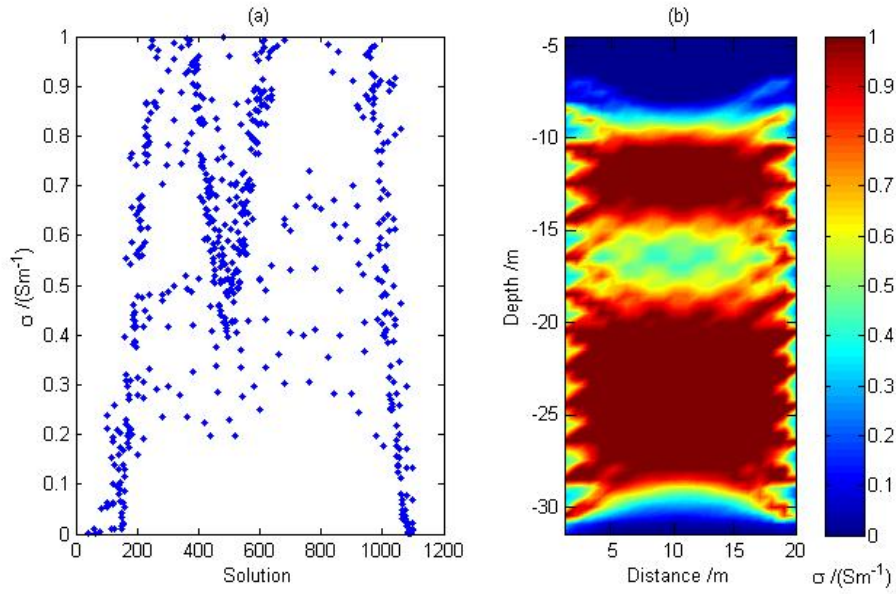


图 6-6 双频电磁波电导率成像 TSVD 反演结果（左）与重建的电导率图像

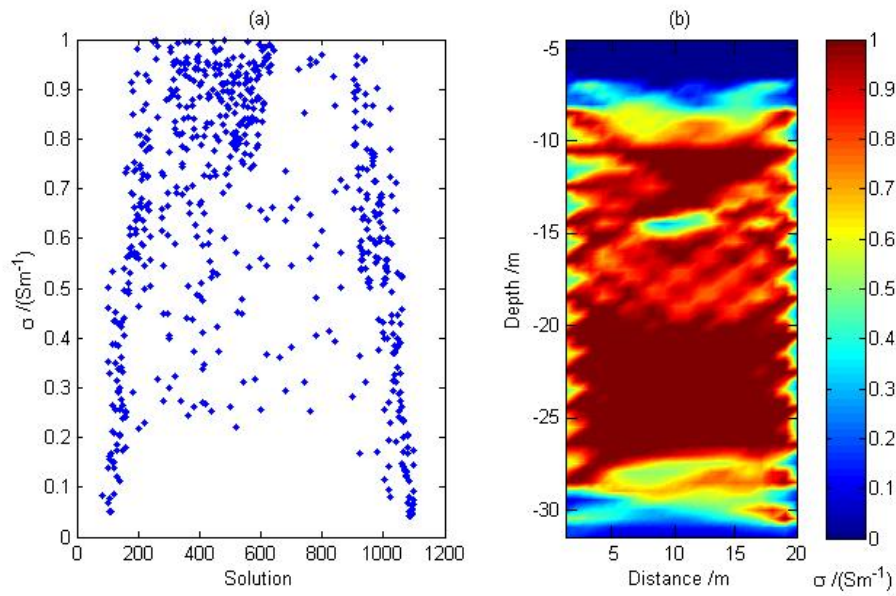


图 6-7 双频电磁波电导率成像 CG 反演结果（左）与重建的电导率图像

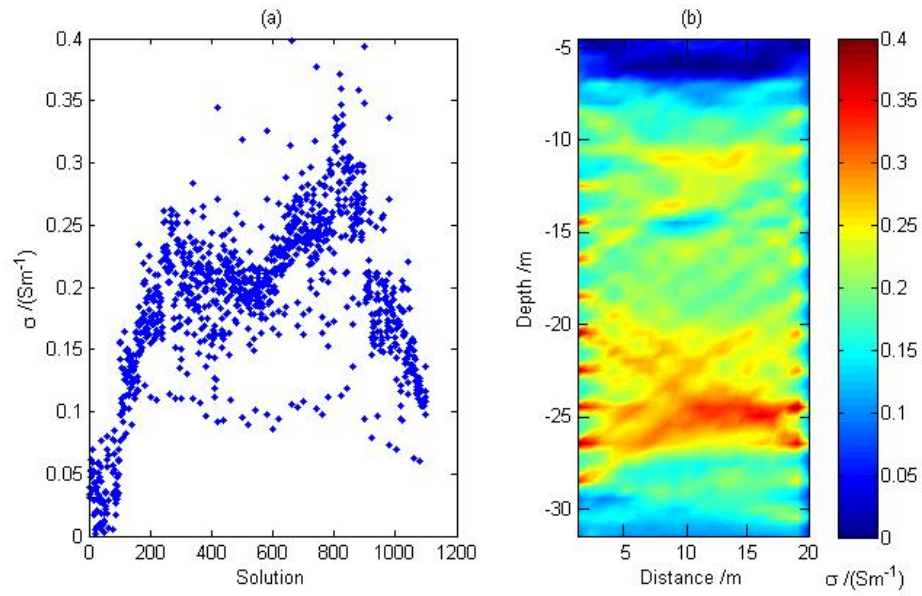


图 6-8 双频电磁波电导率成像标准 Tikhonov 反演结果（左）与重建的电导率图像

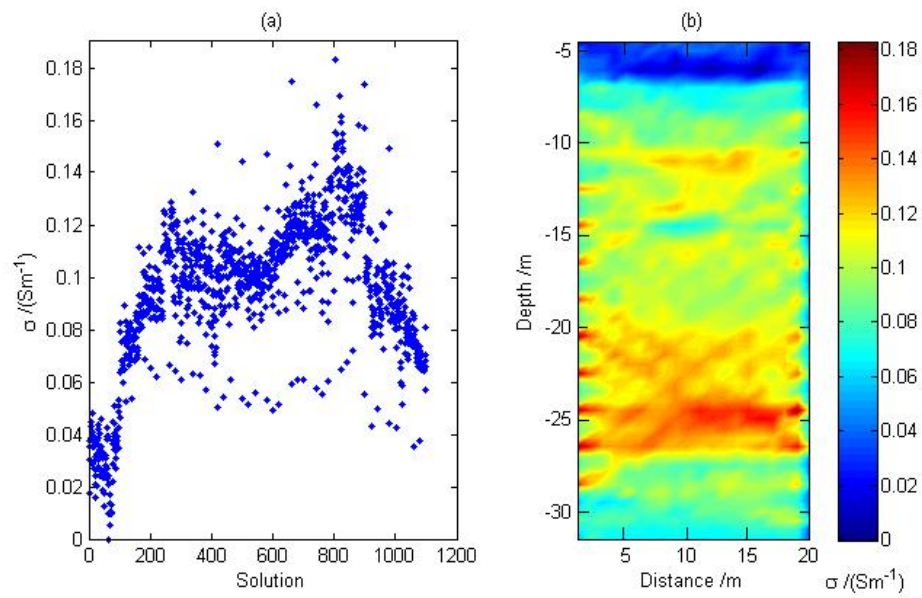


图 6-9 双频电磁波电导率成像双参数混合正则化反演结果（左）与重建的电导率图像

结 论

本文主要针对求解地球物理反演问题时常会涉及到的大型病态矩阵方程的求解问题展开研究. 针对实际反演计算中常会碰到的问题, 将解决此类病态问题最有效的正则化方法与 **Active-Set** 算法、差分进化算法等相结合, 发展了一些新的病态矩阵方程正则化解算方法.

1. 将 **Tikhonov** 正则化方法与 **Active-Set** 算法相结合, 研究在实际问题中, 参数有非负的特性要求时, 将原问题转化为一个带非负约束的阻尼最小二乘问题, 并用 **Active-Set** 算法求解. 通过改进后的迭代算法对理论模型进行了数值模拟计算, 验证了该方法的有效性. 应用到实际电导率成像反演, 取得了较满意的结果. 与常规的 **LSQR**、**SP-LSQR** 方法反演模型比较, 发现不带正则化项的结果明显较差, 与通常利用拟牛顿法进行优化计算的 **Tikhonov** 正则化算法比较发现, 本文方法也有明显的优势, 不管在高噪和低噪反演中, 都能取得稳定的反演结果. 基于 **Tikhonov** 正则化的 **Active-Set** 算法, 克服了大型病态方程的求解困难, 保证了所求的参数为非负的物理限制, 但始终无法避开正则化参数的选取.

2. 进一步研究了差分进化算法在地球物理反演中的应用. 在根据反演的目的与要求, 得到最佳差分进化模型后, 提出了基于种群熵的自适应差分进化(**ARDE**)算法以及粒子群混合差分进化 (**PSODE**)算法与 **Tikhonov** 正则化相结合来进行反演计算. 在不影响反演效果的前提下, 上述两种方法均不同程度地提高了收敛速度, 降低了时间成本. 但要适合于大规模的计算, 还有一定差距. 另外, 本文还提出了一种基于 **LSQR** 的混合差分进化算法. 该算法为一种在迭代正则化的基础上采用 **DE** 进化策略的优化新算法. 该算法具有数值稳定性好、不依赖于初值、不易陷入局部极值和收敛速度快等优点, 尤其是可以避开 **Tikhonov** 和 **TSVD** 等算法的正则化参数的选取困难的情况. 适合于那些对正则化参数敏感性较低, 最优的正则化参数确定困难时的地球物理反演问题.

3. 尝试引入带有二阶正则算子的正则化项, 建立了一种双参数混合正则化方法, 并应用 **L**-曲线法、偏差原理和广义交叉校验准则的优化组合来确定最佳正则化参数, 通过与截断奇异值分解法、共轭梯度法以及拟牛顿法计算的 **Tikhonov** 正则化方法比较, 理论模型的数值模拟结果表明, 在获得较准确的结果的同时, 对于数据的随机扰动具有稳定性. 将此方法应用于电导率的反演计算, 反演成像结果表明了该方法的有效性. 在正则化项增加的情况下, 正则化参数的确定显得更加复杂, 如何寻找更加有效的优化计算方法以确定正则化参数, 将是值得进一步探索的问题.

通过对地球物理反演中病态矩阵方程的各种正则化解法的详细研究,得到了如下结论:本文提出的 A-TR 算法、HDE 算法等方法可针对不同的反演具体情况进行应用,各有其适用情况.如 A-TR 算法适用于反演参数非负要求时的情况,而 HDE 算法适用于正则化参数选取困难时的情况.地球物理反演成像方程通常具有高度的病态性,而每一种反演方法存在一定的局限性,如何在实际情况下,针对不同的情况,寻找适合的正则化算法,与优化方法相结合,设计合理的反演算法,最终得到满意的反演结果,仍然是一个值得深入探索的课题.而通过比较分析,针对本文所着重考察的双频电导率反演问题,由于电导率的非负性,采用 A-TR 方法可得到更加精细可靠的重建图像.用此方法重建的成像图所反映的电性结构特征与钻探所揭露的剖面地质特征更加一致.在成像图的顶部蓝色区域带是低电导率区域,接近地表较近,大概距离地面 4-7 米的地方是干燥松散的泥土和砂砾层;而中下部红色区域带是高电导率区域,深度大概为距离地面 22-28 米,这些地方存在细砂层或高比例细沙和粘土的混合物.同时还可以看出,高导异常体的边界比较清晰,连续性也比较好.

致 谢

本文是在导师曹俊兴教授的悉心指导下完成的。曹教授四年来给予我多方面的指导，我很难在这里悉数，也难以在论文的内容上得到充分的体现。最为重要的是曹教授和我的另一个导师谭永基教授一道，为我打下了必要的理论和实践基础，安排了众多的交叉研究的机会和课题，为我开辟了一个交叉研究的道路。正是由于他们对交叉研究的价值、可能以及迫切性的认识，才让我有机会选择这样一个意义重大但现在回想起来却有相当风险的学业方向，鼓励我实现他们倡导的跨领域、交叉学科研究的总体设想。我深深地感谢他们，并期待能为这个领域的扩展，作继续不懈的努力。

通过我这两位导师的精心指点和关怀，我们的研究受到了牛津大学 John Ockendon 教授（Professor at Oxford University）的充分肯定，并在 Ockendon 教授的邀请和资助下，让我亲自到牛津大学 OCCAM 访问研究。我们的研究结果，不仅受到了 Ockendon 教授的欣赏，并在牛津大学访问期间进一步发展成为和 Ockendon 教授和 Chris Farmer 博士的合作。在这里，我要特别感谢 Ockendon 教授和 Farmer 博士、朱一超博士的热情接待和细致的交流，并希望今后有进一步的更深入的合作。

在我近年来的学习和研究历程中，李大潜院士也给了我很多的指导、鼓励 and 关心。我为自己能得到李教授的启迪和教诲而深感幸运。

感谢所有给我传授专业知识的信息工程学院的老师，特别是贺振华教授、李瑞教授、李录明教授、李才明教授、王绪本教授等，谢谢他们辛勤教育与关怀。

感谢我的合作者潘克家博士，在与他一起进行的学术讨论中，他的意见和建议给了我非常有益的帮助。感谢我的学生冯宝宾同学、谢斌同学。

还有许多其他的亲朋好友，他们给与我的爱护、鼓励、灵感、信心，让我感到温暖和充实。郭朝晖博士、王玮明博士、美国的 Shawn 博士等就是这样的朋友。衷心地感谢他们！其他的，我唯恐有所遗漏，恕不悉数。

本研究得到了“国家自然科学基金 40574032”的资助，在此深表感谢！

谨以此文献给我的父母亲，没有他们的养育教诲，没有他们几十年来一如既往的对我学业上和生活上的全力支持，我无法取得今天的成绩！希望我所取得的成绩可以让他们感到由衷的欣慰。也感谢我可爱的儿子，是他给予了我巨大的精神支持，使我得以顺利完成学业。愿他健康成长！

参考文献

- [1] 王彦飞. 反演问题的计算方法及其应用[M].北京:高等教育出版社,2007
- [2] 何旭初. 数值相关性理论及其应用[J].高等学校计算数学学报, 1979,(1): 11-19
- [3] 朱慈幼. 病态线性方程组的条件预优解法——数值相关性理论的一个应用[J].高等学校计算数学学报,1981,2:97-106
- [4] King J T. On the construction of preconditioners by subspace decomposition[J]. J.of Computational and Applied Math. 1990,29:195-205
- [5] Tikhonov A N. On solving incorrectly posed problems and method of regularization[J]. Dokl. Acad.Nauk USSR,1963,151(3)
- [6] Hadamard J. Lectures on the Cauchy problems in Linear Partial Differential Equations[M].New Haven: Yale University Press,1923
- [7] Tikhonov A N. On stability of inverse problems[J]. Dokl.Acad.Nauk USSR,1943,39(5):195-198
- [8] Phillips D L. A technique for the numerical solution of certain integral equations of the first kind[J]. J.Assoc.Comput.Mach., 1962(9):84-97
- [9] Engl H W. Discrepancy principles for Tikhonov regularization of ill-posed problems leading to optimal convergence rates[J]. J.Optim.Theory.Appl.,1987(52): 209-215
- [10] Engl H W, Gfrerer H. A posteriori parameter choice for general regularization methods for solving ill-posed problems[J]. Appl. Numer. Math..1988(4):395-417
- [11] Engl H W, Kunisch K, Neubauer A. Convergence rates for Tikhonov regularization of non-linear ill-posed problems[J]. Inverse Problems. 1989(5):523-540
- [12] Engl H W. Regularization methods for the stable solution of inverse problems[J]. Surv.Math.Ind.. 1993(3):71-143
- [13] Engl H W, Hanke M, Neubauer A. Regularization of Inverse Problems[M].Dordrecht:Kluwer Academic Publishers,1996
- [14] Hansen P C. Numerical tools for analysis and solution of fredholm integral equations of the first kind[J]. Inverse Problems,1992(8):849-872
- [15] Hansen P C. Analysis of discrete ill-posed problems by means of the L-curve[J]. SIAM Review,1992(34):561-580
- [16] Hanke M, Hansen P C. Regularization methods for large-scale Problems[J]. Surv.Math.Ind., 1993(3):253-315
- [17] Hansen P C. Rank-deficient and Discrete ill-posed Problems:Numerical Aspects of Linear Inversion Philadelphia[M].SIAM Publishers,1998
- [18] Hansen P C. The L_curve and its use in the numerical treatment of inverse problems [J]. Computational Inverse Problems in Electrocardiography. 2001:119-142
- [19] 冯康. 数学物理中的反问题[M].北京:中国科学院计算中心,1985

- [20] 张关泉. 数学物理反演问题[M].北京: 中国科学院计算中心, 1987
- [21] 黄光远, 刘小军. 数学物理反问题[M]. 济南:生动科学技术出版社,1993
- [22] Cheng J, Yamamoto M. One new strategy for a-priori choice of regularizing parameters in Tikhonov's regularization[J]. Inverse Problems,2000,16(4):131-136
- [23] Cheng J, Yamamoto M. Uniqueness in an inverse scattering problem within non-trapping polygonal obstacles with at most two incoming waves[J]. Inverse Problems, 2003 (19), 1361-1384.
- [24] Cheng J, Liu J J, Nakamura G. Recovery of the shape of an obstacle and the boundary impedance from the far-field pattern[J]. Journal of Mathematics of Kyoto University, 2003 (43), 165-186
- [25] Cheng J, Isakov V, Yamamoto M, Zhou Q. Lipschitz stability in the lateral Cauchy problem for elasticity system[J]. Journal of Mathematics of Kyoto University, 2003 (43), 475-501
- [26] Cheng J, Yamamoto M. Global uniqueness in the inverse acoustic scattering problem within polygonal obstacles[J]. Chinese Annals of Mathematics, 2004 (25B), 1-6
- [27] Cheng J, Yamamoto M. Determination of two convection coefficients from Dirichlet to Neumann map in the two-dimensional case[J] SIAM Journal on Mathematical Analysis, 2004 (35), 1371-1393
- [28] Liu J J, Cheng J, Nakamura G. Reconstruction and uniqueness of an inverse scattering problem with impedance boundary[J]. Science in China, Ser.A, 2002,45(11):1408-1419
- [29] Liu J J, Cheng J, Nakamura G. Reconstruction of scattered field from far-field by regularization[J]. Journal of Computational Mathematics, 2004,2(3):389-402.
- [30] Liu J J, On uniqueness and linearization of an inverse electromagnetic scattering problem[J]. Applied Mathematics and Computation.,2005,171(1): 406-419
- [31] 刘继军. 不适定问题的正则化方法及应用[M].北京:科学出版社,2005
- [32] Liu J J, Nakamura G, Potthast R. A new approach and improved error analysis for reconstructing the scattered wave by the point source method[J]. J.Computational mathematics. 2007,25(2)
- [33] Liu J J, Ni M. A model function method for determining the regularizing parameter in potential approach for the recovery of scattered wave[J]. Applied Numerical Mathematics, 2008,58(8):1113-1128
- [34] Sen M K, Roy I G. Two problems in pre-stack inversion: optimal regularization and computation of differential seismograms. SEG 2002 annual meeting, Salt Lake City,USA, 2002
- [35] Roy I G. Iteratively adaptive regularization in inverse modeling with Bayesian outlook-application on geophysical data[J]. Inverse Problems in Science and Engineering. 2005(13):655-670
- [36] Vogel C R. Computational Methods for Inverse Problems. SIAM,2002

- [37] Li X W, Gao F, Wang J D, Strahler A H. A Priori knowledge accumulation and its application to linear BRDF model inversion[J]. Journal of Geophysical Research. 2001(106):11925-11935
- [38] 王彦飞. 数值求解迭代Tikhonov正则化方法的一点注记[J]. 数值计算与计算机应用. 2002,23(3):
- [39] Wang Y F, Wen W Z, Nashed Z, Sun Q Y. A Fast TSVD method Time-limited signal reconstruction,in De-shuang Huang ed.: Advances in Intelligent Computing-ICIC 2005. 2005:51-60
- [40] 栾文贵.地球物理中的反问题[M]. 北京:科学出版社,1989
- [41] 王彦飞, 杨长春, 段秋梁. 地震偏移反演成像的迭代正则化方法研究[J]. 地球物理学报. 2009, 52(6):1615~1624
- [42] 崔岩, 王彦飞, 杨长春. 带先验知识的波阻抗正则化方法研究[J].地球物理学报, 2009, 52 (8) : 2135-2141
- [43] 范宜仁, 魏丹, 邓少贵. Tikhonov正则化方法在测井反演中的应用[J].测井技术, 2005, 29 (2):
- [44] 陈小斌, 赵国泽, 汤吉. 大地电磁自适应正则化反演算法[J].地球物理学报, 2005, 48 (4) :
- [45] 熊小芸, 赵延文, 聂在平. 一种用于电磁逆散射问题的正则化方法[J].自然科学进展, 2006, 10:
- [46] 王家映. 地球物理反演理论[M]. 北京: 高等教育出版社. 2002
- [47] Varah J M. Pitfalls in the numerical solution of linear ill-posed problems[J]. SIAM J.Sci.Stat. Comput.,1983(4):164-176
- [48] 李世雄,刘家琦. 小波变化和反演数学基础. 北京:地质出版社,1994
- [49] 刘劲松,刘福田,刘俊等. 地震层析成像LSQR算法的并行化[J]. 地球物理学报,2006,49(2): 540-545
- [50] Paige, C G, Saunders M A. LSQR: An algorithm for sparse linear equations and sparse least-squares, ACM Trans. Math. Software, 1982, 8:43~71
- [51] 刘劲松,刘福田,刘俊等. 地震层析成像 LSQR 算法的并行化[J].地球物理学报, 2006, 49(2):540-545
- [52] 肖庭延,于慎根,王彦飞. 反问题的数值解法[M]. 北京:科学出版社,2003
- [53] Groetsch C W. The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind. London: Pitman Advanced Publishing Program,1984
- [54] Landweber L. An iteration formula for fredholm integral equations of the first kind[J].Amer. J.Math.,1951(73):615-624
- [55] Fridman V. Methods of successive approximation for fredholm integral equations of the first

- kind[J].Uspekhi Mat.Nauk,1956,11(1):233-234
- [56] Kunisch K. On a class of damped Morozov principle[J]. Computing,1993(50):185-198
- [57] Kunisch K, Zou J. Iterative choices of regularization parameter in linear inverse problems[J]. Inverse Problems,1998(14):1247-1264
- [58] Neumaier A. Solving ill-conditioned and singular linear systems:A tutorial on regularization[J]. SIAM Rev.,1998,40(3):636-666
- [59] Vainikko G M. The discrepancy principle for a class of regularization methods[J]. USSR Compt. Math.Phys.,1983(23):1-6
- [60] Vainikko G M, Varenennikov A J. Iteration procedures in ill-posed problems[J]. Moscow : Nauka,1986
- [61] Varah J M. On the numerical solution of ill-conditioned linear systems with application to ill-posed problems[J]. SIAM J.Numer.Anal.,1973(10):257-267
- [62] Vogel C R. Non-convergence of the L-curve regularization parameter selection method[J]. Inverse Problems,1996(12):535-547
- [63] Xiao T Y. On the algorithms for solving discrepancy equation in ill-posed problems. '96 Symposium on computation physics for Chinese Overseas and at Home,1996,Beijing
- [64] Morozov V A. Methods for Solving Incorrectly Posed Problems. New York:Springer,1984
- [65] Morozov V A. On regularization of ill-posed problems and selection of regularization parameter[J]. J.Comp.Math.Math.Phy.,1966,6(1):170-175
- [66] Wang Y F, Xiao T Y. Fast realization algorithms for determining regularization parameters in linear inverse problems[J].Inverse Problems,2001(17):281-291
- [67] Groetsch C W. The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind. Pitman Advanced Publishing Program,1984
- [68] Chen H, Hou Z Y. The iterated regularization with perturbed operators and noisy data[J]. Science in China(series A),1994(37):1409-1417
- [69] Tikhonov A N, Arsenin V Y. Solutions of Ill-posed Problems. New York: John Wiley and Sons, 1977
- [70] Golub G H, Heath M, Wahba G. Generalized cross-validation as a method for choosing a good ridge parameter[J]. Technometrics,1979(21):215-223
- [71] Eldén L. A note on the computation of the generalized cross-validation function for ill-conditioned least squares problems[J]. BIT,1984(24):467-472
- [72] 曹俊兴, 聂在平. 线性方程组迭代解的随机模型测试研究[J]. 物探化探探测技术. 1998, 20(2): 120-124
- [73] 张宏兵, 尚作萍, 杨长春, 段秋梁. 波阻抗反演正则参数估计[J]. 地球物理学报, 2005, 48(1): 181-188
- [74] Wang Y F, Xiao T Y. Fast convergence algorithms for solving 2-D integral equations of the first kind, In chen et al. Lecture Notes in Physics[J]. Springer-Verlag, 2000: 333-344
- [75] Wang Y F. A restarted conjugate gradient method for ill-posed problems[J]. Acta

- Mathematica Applicata Sinica. Springer-Verlag,2003(19):31-40
- [76] Wang Y F. Iterative Choices of Regularization Parameter with Perturbed Operator and Noisy Data. Report,AMSS,Chinese Academy of Sciences,2001
- [77] 曹俊兴. 双频电磁波电导率层析成像反演[J]. 地球物理学报,2001, 44(增刊): 199-205
- [78] Cao J X, He Z H, Zhu J S, et al. Conductivity Tomography at two frequencies[J]. Geophysics,2003, 68(2):516-522
- [79] 王文娟,潘克家,曹俊兴,谭永基.基于 Tikhonov 正则化的双频电磁波电导率反演成像[J].地球物理学报.2009, 52(3):750-757
- [80] Wang W J, Pan K J, Cao J X, Tan Y J. Conductivity Tomography Imaging using Regularization Methods[J]. Journal of Information and Computation Science. 2008,5(4): 1885-1891
- [81] Feng B B, Wang W J, Cao J X. An Improved Tikhonov Regularization Method for Conductivity Tomography Imaging. 2008 International Symposium on Computer Science and Computational Technology (ISCSCCT2008) . 2008,Volume 2:406 - 409.
- [82] Golub G H, van Loan C F. Matrix Computations[M].Johns Hopkins Univ. Press, 1989. Second Edition
- [83] 曹俊兴,聂在平. 基于物理模型的成像算法评价与 SASART 成像算法[J].成都理工学院学报,1998,25(4):473-479
- [84] 杨文采,杜剑渊. 层析成像新算法及其在工程检测上的应用[J].地球物理学报,1994,37(2): 293-244
- [85] 常旭,卢孟夏,刘伊克. 地震层析成像反演中 3 种广义解的误差分析与评价[J].地球物理学报,1999,42(5):695-701
- [86] 刘伊克,常旭. 地震层析成像反演中解的定量评价及其应用[J].地球物理学报, 2000, 43(2): 251-256
- [87] 王振宇,刘国华,梁国钱. 基于广义逆的层析成像反演方法研究[J].浙江大学学报(工学版), 2005,39(1):1-5
- [88] 顾勇为,归庆明,边少锋等. 地球物理反问题中两种正则化方法的比较[J].武汉大学学报(信息科学版), 2005,30(3):238-241
- [89] 刘玉柱,董良国,夏建军. 初至波走时层析成像中的正则化方法[J].石油地球物理勘探,2007, 42(6): 682-685
- [90] Bjorck A. A direct method for sparse least-squares problems with lower and upper bounds[J]. Numer. Math., 1988(54):19-32
- [91] Lotstedt P. Perturbation bounds for the linear least-squares problem subject to linear inequality constrains[J].BIT,1983,23:500-519
- [92] Murty K. Note on a Bard-type scheme for soving the complementarity problem[J].Opsearch,1974,11:123-130

- [93] Lawson C L, Hanson R J. Solving Least-Squares Problems[M].Englewood Cliffs: Prentice-Hall, 1974
- [94] 杨文采. 非线性地球物理反演方法:回顾与展望[J].地球物理学进展, 2002, 17(2):255-261
- [95] Ceperley D, Alder B. Quantum Monte Carlo[J].Science, 1986, 231(4738):555-560
- [96] 魏超, 李小凡, 张美根. 基于量子蒙特卡罗的地球物理反演方法[J].地球物理学报, 2008, 51(5):1494-1502
- [97] Stoffa P L, Sen M K. Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane-wave seismograms[J].Geophysics 1991, 56:1794-1810
- [98] 师学明,王家映,张胜业等. 多尺度逐次逼近遗传算法反演大地电磁资料[J].地球物理学报, 2000,43 (1): 122-130
- [99] 罗红明, 王家映, 朱培民等. 量子遗传算法在大地电磁反演中的应用[J].地球物理学报, 2009, 52(1): 260-267
- [100] Kirkpatrick S, Gelatt C D, Vecchi M P. Optimization by simulated annealing[J].Science, 1983, 220 (4598) : 671-680
- [101] 姚姚. 地球物理非线性反演模拟退火法的改进[J].地球物理学报, 1995, 38(5): 643-650
- [102] 魏超, 朱培民, 王家映. 量子退火反演的原理和实现[J].地球物理学报, 2006, 49(2): 577-583
- [103] 师学明,王家映,易远元等. 一种新的地球物理反演方法——模拟原子跃迁反演法[J].地球物理学报,2007 ,50 (1) :305-312
- [104] 潘克家,王文娟,谭永基,曹俊兴. 基于混合差分进化算法的地球物理线性反演[J].地球物理学报, 2009,52(12):3083-3090
- [105] 王文娟, 谢斌, 潘克家等. 利用加速差分进化算法反演非均匀介质电磁成像[J].地球物理学进展(已录用)
- [106] Storn R, Price K. Differential Evolution — A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces[J]. J. Glob. Optim. 1997, 11:341-359
- [107] 阳明盛,罗长童. 最优化原理、方法及求解软件[M].北京:科学出版社, 2006
- [108] 潘克家,陈华,谭永基.基于差分进化算法的核磁共振 T_2 谱多指数反演[J].物理学报, 2008, 57(9):5956-5961
- [109] 韩伟,李道伦,卢德堂,陈锋.利用改进差分进化算法的油田地层参数反演[J].电子技术, 2008,8:74-78
- [110] Harry Eckel, aus Kassel. Numerical study of an evolutionary algorithm for electrical impedance tomography . der Georg-August Universitat Gottingen, 2007
- [111] Holland J. H. Genetic Algorithm[J]. Science.1992, 11:24-31
- [112] 陈良,戴光明,张全元,谢柏桥.差分演化算法及其改进形式的综述[J].计算机工程与设计.2008, 29(1):131-135

- [113] Gamperle R, Dmuller S, Koumoutsakos P. A parameter study for differential evolution [A]. International conference on advances in intelligent systems, Fuzzy systems, Evolutionary Computation[C]. 2002:293-298
- [114] 谢晓峰,张文俊,张国瑞等.差分演化算法的实验研究[J].控制与决策,2004,19(1):45-52
- [115] Godfrey C.Onwubolu, Donald Davendra. Differential Evolution: A Handbook for Global Permutation-Based Combinatorial Optimization [M].Springer, 2008
- [116] Hansen P C. The Truncated SVD as a Method for Regularization. BIT, 1987,27:534-553
- [117] Josef Tvrdik. Adaptation in differential evolution: A numerical comparison [J]. Applied Soft Computing,2009,9:1149-1155
- [118] 邓泽喜,曹敦虔,刘晓冀,李娜.一种新的差分进化算法[J].计算机工程与应用.2008, 44(24):40-44
- [119] 吴亮红,王耀南,袁小芳,周少武.自适应二次变异差分进化算法[J].控制与决策.2006, 21(8):898-902
- [120] 江瑞,罗予频,胡东成,司徒国业.一种基于种群熵估计的自适应遗传算法[J].清华大学学报.2002, 42(3):358-361
- [121] 宋立明,李军,丰镇平.ARDE 算法及其在三维叶栅气动优化设计中的应用[J].工程热物理学报.2005, 26(2):221-224.
- [122] Kennedy J, Eberhart R C. Particle Swarm Optimization. Proc. IEEE International Conference on Neural Networks, Perth, Australia, 1995:1942-1948
- [123] Kennedy J, Eberhart R C. A Discrete Binary Version of the Particle Swarrm Algorithm. Proc. 1997 Conference on Systems, Man and Cybernetics. 1997:4104-4109
- [124] Angeline P J. Using Selection to Improve Particle Swarm Optimization. Proc. IEEE International Conference on Evolutionary Computation. Alaska, USA,1998:84-89
- [125] Angeline P J. Evolutionary Optimization Versus Particle Swarm Optimization:Philosophy and Performance Differences. The Seventh Annual Conference on Evolutionary Programming. 1998
- [126] Lovbjerg M, Ruamussen T K, Krink T. Hybrid Particle Swarm Optimizer with Breeding and Subpopulations. Proc. The Third Genetic and Evolutionary Computation Conference, 2001
- [127] Kennedy J, Eberhart R C, Shi Y. Swarm Intelligence. San Francisco: Morgan Kaufmann Publishers, 2001.
- [128] 高尚,杨静宇.群智能算法及其应用[M].北京:中国水利水电出版社.2006 年
- [129] 贺安坤,苗良.差分进化微粒群优化算法——DEPSO[J].微计算机信息.2006, 22(3):284-276.
- [130] 栾丽君,谭立静,牛奔.一种基于粒子群算法和差分进化算法的新兴混合全局优化算法[J].信息与控制.2007, 36(6):708-714
- [131] Aliev B 1986 *Vestnik Moskovskogo Universiteta, Vychislitel'naya Matematikai*

- Kibernetika*. 2 45
- [132] Aliev B 1991 *Vestnik Moskovskogo Universiteta, Vychislitel'naya Matematikai Kibernetika*. 1 28
- [133] 杨宏奇,侯宗义. 非线性不适定算子方程的双参数正则化方法[J].数学物理学报, 2004, 24A (2) : 129-134

附录： 博士研究生学习期间发表和完成的论著

- [1]. 王文娟, 潘克家, 曹俊兴等. 基于 Tikhonov 正则化的双频电磁波电导率层析成像. 《地球物理学报》(SCI 收录) 2009, 52 (3): 750-757;
- [2]. 潘克家, 王文娟, 等. 基于混合差分进化算法的地球物理线性反演. 《地球物理学报》(SCI 收录) 2009, 52 (12): 3083-3090;
- [3]. Lei Zhang , Wenjuan Wang , Yakui Xue. Spatiotemporal complexity of a predator-prey system with constant harvest rate. Chaos, Solitons & Fractals (SCI、EI 收录) 2009, 41(1):38-46;
- [4]. Wenjuan Wang, Bin Xie, Yusheng Li, Kejia Pan. The Evaluation and Application Research about Regional Innovation Capability Based on Rough Set and BP Neural Network. 2009 Second International Conference on Modelling and Simulation (ICMS2009) (EI 收录),2009, Volume 3:308 - 311;
- [5]. Wenjuan Wang, Kejia Pan, Junxing Cao. Conductivity Tomography Imaging using Regularization Methods. 《Journal of Information and Computation Science》 (EI 收录). 2008,5(4): 1885-1891;
- [6]. Baobin Feng, Wenjuan Wang (Corresponding author), Junxing Cao. An Improved Tikhonov Regularization Method for Conductivity Tomography Imaging . 2008 International Symposium on Computer Science and Computational Technology (ISCST2008) . (EI 收录)2008,Volume 2:406 – 409;
- [7]. 王文娟, 曹俊兴, 等. 基于微粒群算法的神经网络储层物性参数预测. 《西南石油大学学报》(EI 收录) 2007, 29(6): 31-33;
- [8]. 王文娟, 张雷. 带有控制结构约束的次优状态反馈律的研究. 《成都理工大学学报》(自然科学版) 2007, 34 (6): 697-700;
- [9]. 陈克勇, 王文娟, 等. 长岭气田营城组火山岩储层含水饱和度建模. 《石油天然气学报》(中文核心期刊) 2009, 31 (4): 77-80;
- [10]. 张章华, 孟凡卉, 王文娟. 基于微粒群算法的神经网络在岩性识别中的应用. 《物探化探计算技术》2007, 29(2): 130-133;
- [11]. 王文娟, 谢斌, 等. 利用加速差分进化算法反演非均匀介质电磁成像. 《地球物理学进展》(已录用);
- [12]. 王文娟, Chris Farmer 等. 双参数混合正则化方法及在地球物理反演中的应用. (已投稿).