

求解非线性反问题的大范围收敛梯度正则化算法

崔凯¹, 李兴斯^{*2}, 李宝元², 杨国伟¹

(1 中国科学院力学研究所 高温气体动力学重点实验室, 北京 100080;

2 大连理工大学 工程力学系, 辽宁 大连 116023)

摘要: 基于同伦映射的思想, 改进了求解非线性反问题的梯度正则化算法。通过路径跟踪有效地拓宽了梯度正则化算法求解的收敛范围。对于正则化参数的修正, 通过引入拟 Sigmoid 函数, 提出了一种下降速率可调的连续化参数修正方法, 在保证迭代稳定的条件下, 得到较好的计算效率, 同时保证该算法具有很好的抵抗观测噪声能力。实际算例表明, 该方法收敛范围宽, 计算效率高, 在存在较强观测噪声的条件下也能得到很好的反演结果。

关键词: 反演; 梯度正则化; 同伦方法; 正则化参数

中图分类号: O 39; TB 12 **文献标识码:** A

1 引言

非线性反问题广泛地存在于许多科学和工程问题中。反问题求解的主要困难在于问题的不适定性, 即待求参量不连续依赖于带有误差的观测数据。正则化方法^[1,2]是求解反问题的一种有效方法, 该方法通过引入正则化参数及镇定泛函将目标泛函连续化, 从而得到解的一个稳定近似, 即正则解。

迭代正则化方法^[2-4]是近年来出现的求解非线性反问题的一种有效方法。该方法通过求解控制方程的 Frechet 导数, 将非线性问题转化为一系列线性问题求解, 其正则化效应通过提前终止迭代体现, 从而达到抵抗观测噪声的目的。迭代正则化方法可分为显式方法(如最速下降法、Landweber 迭代方法等)和隐式方法(如牛顿型方法、拟牛顿法等)两类。显式方法所需要的迭代步数相对较多, 而隐式方法虽然求解迭代步数要少于显式方法, 但每一迭代步中的计算量则要多于显式方法^[5]。

梯度正则化算法(GR)^[6]属于显式迭代正则化方法。自提出以来, 在声波方程反演^[7]、材料物性参数识别^[8]等许多领域都得到了成功的应用。较人工神经网络、遗传算法和模拟退火算法等启发式算法而言, 该方法具有较高的计算效率。同伦方法^[9]是一种大范围收敛方法。该方法通过构造同伦映射, 延拓求解路径, 可以有效地扩大收敛域, 降低求解

过程对于迭代初值的依赖性。本文基于同伦方法的思想, 对于梯度正则化算法进行修正。构造出一种大范围收敛梯度正则化算法(HGR), 在保持梯度正则化算法高效稳定的条件下, 进一步拓宽了梯度正则化算法的收敛域。此外通过引入拟 sigmoid 函数, 提出了一种步长可调的连续化正则化参数修正方法, 有效地避免了在迭代过程中由于试探正则化参数而带来的多余正演计算, 进一步提高了计算效率, 并同时保证其具有较好的抵抗观测噪声能力。

2 大范围收敛梯度正则化算法求解迭代公式

对于一般科学和工程问题而言, 其数学模型一般可以表达为微分方程的形式, 即

$$L[p(x), u(x, t)] = 0 \quad (1)$$

其中 L 为方程算子, $p(x)$ 为方程参量, $u(x, t)$ 为方程的待求量。当参量 $p(x)$ 的值已知时, 通过数值求解可以计算得到问题的解 $u^c(x, t)$ 。此过程称为正演。若给定不同参量 $p(x)$, 通过正演总能得到相应的解 $u(x, t)$, 即两者之间存在着一种映射关系 G , 使得:

$$G[p(x)] = u(x, t) \quad (2)$$

对于某些实际问题而言, 参量 $p(x)$ 很难直接测得, 因此必须通过某些边界上的观测值 $u^*(x, t)$ 根据方程(1)反演参量 $p(x)$ 的值。这就构成了一个反问题。若算子 G 为非线性算子, 则为非线性反问题。对于实际问题而言, 观测值 $u^*(x, t)$ 总是带有观测噪声, 即

收稿日期: 2003-07-18; 修改稿收到日期: 2004-06-28.

基金项目: 国家重点基础研究发展规划(G1999032805)资助项目.

作者简介: 崔凯(1973-), 男, 博士后;

李兴斯^{*}(1942-), 男, 博士生导师.

$$u^*(x, t) = u^T(x, t) + \eta(x, t) \quad (3)$$

其中 $u^T(x, t)$ 为不含观测噪声的真观测值, $\eta(x, t)$ 为观测噪声。此时, 问题(2) 将不能精确满足。在这种情况下, 可以将参量 $p^*(x)$ 的求解转化为如下最优化问题, 求得该问题的最优解, 即

$$\min J(p(x)) = \int_{\Omega} G[p(x)] - u^*(x, t) \, d\Omega \quad (4)$$

式(4) 即为非线性反问题的数学模型。其中 α 为某个范数, 一般选取为 2 范数。

一般来说, 式(4) 为不适定的。对于这类不适定问题, Tikhonov 正则化^[11] 是一种有效的求解方法。取式(4) 中范数为 2 范数, 则问题(4) 的 Tikhonov 正则解为

$$p_{\alpha}(x) = \inf_{p(x)} \{ J_{\alpha}(p(x)) \} = \inf_{p(x)} \left\{ \int_{\Omega} G[p(x)] - u^*(x, t) \, d\Omega + \alpha \|p(x) - p^0(x)\|^2 \right\} \quad (5)$$

其中 α 为正则化参数, $\|p(x) - p^0(x)\|^2$ 为正则化泛函, $p^0(x)$ 为求解初值。将 $G[p(x)]$ 在 $p^0(x)$ 处进行一阶 Taylor 展开, 有

$$G[p(x)] = u^0(x, t) + \int_{\Omega} \frac{\partial G}{\partial p} \Big|_{p=p^0} \delta p \, d\Omega \quad (6)$$

将式(6) 代入式(5), 并取目标泛函变分为零, 整理得

$$\int_{\Omega} F_{p^0} [\int_{\Omega} F_{p^0} \delta p \, d\Omega] d\Omega + \alpha \delta p = \int_{\Omega} F_{p^0} (u^0(x, t) - u^*(x, t)) \, d\Omega \quad (7)$$

其中 F_{p^0} 为算子 G 在 p^0 处的 Frechet 导数。将式(7) 离散化, 即可得到如下的梯度正则化方法的迭代公式:

$$\begin{cases} [\{K^{(n)}\}^T \{K^{(n)}\} + \alpha I] d\tilde{p} = \{K^{(n)}\}^T (\tilde{u}^* - \tilde{u}^{(n)}) \\ \tilde{p}^{(n+1)} = \tilde{p}^{(n)} + d\tilde{p} \\ \tilde{u}^{(n+1)} = G(\tilde{p}^{(n+1)}) \end{cases} \quad (8)$$

式中 矩阵 $\{K^{(n)}\}$ 为第 n 个迭代步时离散后的导数矩阵; $\tilde{u}^{(n)}$ 为第 n 个迭代步的计算值矢量。

同伦算法^[9] 是一种大范围收敛算法, 该方法通过构造并求解原方程的同伦映射, 可以有效的扩大原问题的求解收敛域。若需要求解的原方程为

$$F(x) = 0 \quad (9)$$

应用同伦算法求解问题(9), 可构造原问题的不动点同伦函数:

$$H(x, \lambda) = (1 - \lambda)F(x) + \lambda(x - x_0) \quad (10)$$

其中 λ 称为同伦参数。显然, 当 λ 为 0 时, 问题(10) 完全对应于原问题(9), 当 λ 为 1 时, 则问题(10) 对应于一个简单的平凡映射。若从问题(10) 出发, 在求解过程中按照某种方式不断减小 λ 的数值, 最终总可以通过一个合理而稳定的求解路径得到原问题(9) 的解。通过这种路径跟踪, 可以有效地扩大原问题的计算收敛域。若依据同伦算法的思想, 对梯度正则化法加以修正, 则可以在保持梯度正则化法原有优点的条件下, 进一步扩大计算收敛域。

根据上述思路, 将式(5) 做不动点同伦修正为

$$p_H(x) = \inf_{p(x)} \left\{ \int_{\Omega} (1 - \lambda) G[p(x)] - u^*(x, t) \, d\Omega + \lambda \|p(x) - p^0(x)\|^2 \right\} \quad (11)$$

仿照梯度正则化法的推导过程, 最后可导出大范围收敛梯度正则化法(HGR) 的迭代公式为

$$\begin{cases} [(1 - \lambda^{(n)}) \{K^{(n)}\}^T \{K^{(n)}\} + \lambda^{(n)} I] d\tilde{p} = \\ (1 - \lambda^{(n)}) \{K^{(n)}\}^T (\tilde{u}^* - \tilde{u}^{(n)}) \\ \tilde{p}^{(n+1)} = \tilde{p}^{(n)} + d\tilde{p}, \quad \tilde{u}^{(n+1)} = G(\tilde{p}^{(n+1)}) \end{cases} \quad (12)$$

式中 $\lambda^{(n)}$ 为同伦参数, 对比式(8) 和式(12), 可以发现式(8) 中的正则化参数与式(12) 中的同伦参数有着十分明显的对应关系。因此, 式(12) 中的同伦参数等价于式(8) 中的正则化参数。

3 正则化参数(同伦参数)的选取和算法的实现

3.1 正则化参数(同伦参数)的选取

在梯度正则化法中, 如何适当地选择正则化参数是一个非常关键的问题。首先, 正则化参数的变化过程应是一个由大到小的渐变过程, 以逐步抵抗反问题的不适定性。在计算的初始阶段, 该参数应该选择较大, 以保证求解的稳定进行, 避免由于问题的病态而引起的迭代不稳定。而随着迭代的进行, 该参数应逐渐减小, 以使得问题(5) 不断地逼近原问题(4)。其次, 因为实际问题必然包含一定的观测噪声, 通过矩阵的奇异值分解可以知道, 正则化参数的数值最终必然不能为零, 其大小与观测噪声的强度有关。

文献[8] 对于正则化参数的选取作了探讨, 其方法借鉴了信赖域法的思路。主要思路为: 先赋予正则化参数一个比较大的值, 以使得问题(8) 的求解是适定的。在迭代过程中, 按照某个比例不断减小该参数, 若在某迭代步中目标函数没有减小, 则

再按照某个比例增大该参数。直到迭代结束。这种正则化参数的选择方法是一种间断型的选择方法, 其中有两个问题需要注意: 首先, 正则化参数的初值确定需要凭借经验, 若该参数选择较小, 则不能保证迭代稳定, 若选择较大, 虽可以保证计算稳定, 但会相应地增加计算量, 带来一些不必要的计算过程; 其次, 在迭代过程中, 由于正则化参数的逆向调整会带来一些不必要的多余正演计算过程, 从而增加计算消耗。

对于式(12) 给出的大范围收敛梯度正则化算法, 正则化参数的选取也应相应遵循以下两个原则: (1) 正则化参数(即同伦参数) 被界定与 [0, 1] 之间, 其初值最大值为 1, 这样正则化参数的初值可以选取为 1 或小于 1 的数值, 在迭代中按照某种方式稳定的下降。(2) 正则化参数的最终取值为接近于零的数值, 其最终的具体数值与观测误差相关, 一般为接近于零的数值。

根据以上两条原则, 我们参考人工神经网络的传递函数, 即 Sigmoid 函数(或称 S 函数) 提出如下的连续化正则化参数的调整公式(拟 Sigmoid 函数, 如图 1 所示)

$$\lambda(N) = \frac{1}{1 + e^{-\beta(N - N_0)}} \quad (13)$$

其中 N 为迭代步数, $\lambda(N)$ 为第 N 个迭代步时正则化参数的取值, β 为下降速率参数, N_0 为初值选取参数。该函数的函数值界于 [0, 1] 之间, 并且在正无穷大方向趋近于 0, 而不等于 0。此外, 该函数是光滑函数, 这种连续的下降过程将更加利于求解的稳定性, 减小正演次数。

这里有两个可调参数, 即 β 和 N_0 。其中 β 用于调整正则化参数的下降速率。该参数取值越大, 则正则化参数的下降速度越快, 相应的迭代收敛速度也快, 但若该参数取值过大, 则可能引起迭代的发散, 一般建议该参数取为 0.5 左右的数值, 这样可以兼顾计算效率和稳定性。参数 N_0 为正则化参数

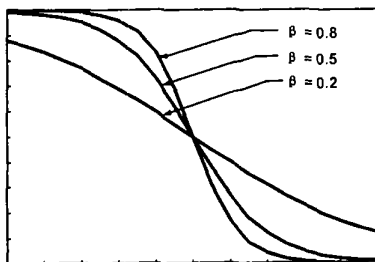


图 1 不同 β 下的拟 Sigmoid 函数曲线
Fig 1 Quasi-sigmoid function with different β

从迭代初值下降至 0.5 的预估迭代步数, 一般可以选为 0 ~ 5 之间的整数, 主要用于保证计算初期的迭代稳定进行。

3.2 算法实现

综合前面的讨论, 可以总结出求解非线性反问题的大范围收敛梯度正则化(HGR) 算法如下:

- step 1: 给定待反演参量的初始值 $\tilde{p}^{(0)}$, β 和 N_0 的初值, 迭代终止误差 ϵ , 置迭代步 $k = 0$;
- step 2: 根据正问题控制方程通过数值计算求得计算结果 $G(\tilde{p}^{(k)})$, Jacobian 矩阵 $K^{(k)}$, 若满足 $G(\tilde{p}^{(k)}) - \tilde{u}^* < \epsilon$ 则终止计算, 此时得到的参量矢量 $\tilde{p}^{(k)}$ 即为反演得到的正则解, 否则继续;
- step 3: 根据(13) 求得该迭代步的正则化参数数值 $\lambda^{(k)}$;
- step 4: 求解方程(12), 得到参量调整向量 $\tilde{\Delta p}^{(k)}$;
- step 5: $\tilde{p}^{(k+1)} = \tilde{p}^{(k)} + \tilde{\Delta p}^{(k)}$, 并置 $k = k + 1$, 转 step 2。

4 数值算例

非饱和土中带有非平衡吸附的溶质迁移过程可由如下方程描述:

$$\begin{cases} K_{cd}\rho \frac{\partial C}{\partial t} + \frac{\partial(\theta C)}{\partial t} + \rho K_A \gamma(1 - S_2)C - \rho K_B S_2 = \\ \frac{\partial}{\partial z} \left(\theta v \frac{\partial C}{\partial z} \right) - \frac{\partial(qC)}{\partial z} \\ \frac{\partial S_2}{\partial t} = K_A(1 - S_2)C - K_B S_2 \end{cases} \quad (14)$$

式中 ρ 为土壤密度 (cm/g^3), θ 为饱和度(体积含水率), α 为纵向弥散度, v 为液相流速 (cm/h), q 为达西流速 (cm/h), K_{cd} 为吸附系数, K_A 为沉淀速度, K_B 为解吸速度, γ 为相对最大沉淀量, 其中未给出量纲之参数均为无量纲参数; C 为污染物相对浓度, S_2 为相对吸附浓度。

在实际问题中, α , K_{cd} , K_A , K_B 和 γ 不能通过实验直接测得。可以通过土柱试验, 根据观测得到的土柱边界溶质浓度值来反演以上参数, 其余参数可直接测得。

对于这个问题, 我们从不同的角度在相同参数下分别使用梯度正则化法(GR) 和本文提出的大范围收敛梯度正则化法(HGR) 进行反演计算。从收敛域、计算效率和抵抗噪声能力三个方面进行了比较。

表 1 不同初始值条件下的反演结果

Tab 1 Computation results of inversion with different initial values

初始解					反演	反演结果					迭代
α	K_{Cd}	K_A	K_B	γ	方法	α	K_{Cd}	K_A	K_B	γ	次数
5.0	0.5	5.0	0.5	0.5	GR	计算发散					—
					HGR	0.6849	0.1149	1.9360	0.0025	0.6037	19

表 2 算例 2 的反演计算结果

Tab 2 Results of example 2

初始解					反演	反演结果					迭代
α	K_{Cd}	K_A	K_B	γ	方法	α	K_{Cd}	K_A	K_B	γ	次数
0.5	0.5	3.0	0.5	0.5	GR	0.6764	0.1151	1.9151	0.0026	0.6035	24
					HGR	0.6764	0.1151	1.9151	0.0026	0.6035	16

数值模拟的基本过程为: 在已知可直接测得参量的条件下, 首先人为给定一组待反演参数数值, 这里我们给定 $[\alpha K_{Cd} K_A K_B \gamma] = [0.6835 \ 0.1154 \ 1.9298 \ 0.00255 \ 0.6037]$, 之后使用有限元法求解方程(14), 计算得到边界溶质浓度值的精确结果, 在此结果上加入一定强度的噪声, 以模拟实际的观测噪声。最后以加入噪声后的计算结果作为反演根据, 给定迭代中止误差和反演初始值, 分别采用以上两种算法进行反演。反演结束后将反演后得到的参数与事先给定的参数相比较以评价反演方法。

算例 1 计算收敛域的比较

以 $[\alpha K_{Cd} K_A K_B \gamma] = [0.5 \ 0.5 \ 5.0 \ 0.5 \ 0.5]$ 为反演起始值, 使用两种算法进行反演。本算例中, 加入的相对噪声强度为 0.1%, 迭代中止误差为 10^{-4} , 式(13)中的两个调整参数分别取为: $\beta = 0.5, N_0 = 0$ 。计算结果如表 1 所示。

算例 2 计算效率的比较

以 $[\alpha K_{Cd} K_A K_B \gamma] = [0.5 \ 0.5 \ 3.0 \ 0.5 \ 0.5]$ 为反演起始值, 采用两种算法进行反演计算。本算例中, 加入的相对噪声强度为 0.1%, 噪声分布由计算机随机产生, 与算例 1 的噪声分布不同。迭代中止误差为 10^{-4} , 式(13)中的两个参数分别为: $\beta = 0.5, N_0 = 0$ 。计算结果如表 2 所示。

算例 3 不同噪声强度下的计算结果

采用与算例 1 相同的迭代中止误差, 以 $[\alpha K_{Cd} K_A K_B \gamma] = [0.5 \ 0.5 \ 3.0 \ 0.5 \ 0.5]$ 为反演起始值, 在不同噪声强度下采用两种算法进行反演计算。式(13)中的两个参数分别为: $\beta = 0.5, N_0 = 0$ 。反演误差采用两种计算方式, 分别为所有参数同真实参数间的 2 范数误差以及参数间的最大误差。计

算结果如表 3 所示。

表 3 不同噪声强度下的反演计算结果

Tab 3 Results of different noise level

噪声强度	反演方法	反演精度		迭代次数
		2 范数误差	最大误差	
1%	GR	0.0316	0.0095	35
	HGR	0.0316	0.0095	18
2%	GR	0.0604	0.0193	25
	HGR	0.0604	0.0193	19
5%	GR	0.1506	0.0462	32
	HGR	0.1506	0.0461	18
8%	GR	0.1996	0.0821	36
	HGR	0.1996	0.0821	19

5 讨论和结论

从算例 1 的计算结果可以看出, 本文提出的大范围收敛梯度正则化算法有效的改进了原算法的收敛域。在存在一定观测误差的条件下, 即使计算初始值选取距离解较远, 在梯度正则化算法计算发散的情况下, 采用大范围收敛梯度正则化算法计算仍能够收敛到一个很好的解点。这是由于对梯度正则化法进行同伦修正后, 通过路径跟踪^[9, 10], 将同伦算法的大范围收敛性引入了梯度正则化法, 从而增大了原算法的收敛域。从算例 2 的计算结果可以看出, 在相同的计算条件下, 梯度正则化法同本文提出的方法都能得到很好的反演计算结果, 但采用改进后的方法其计算效率明显高于梯度正则化法。这是由于采用了连续化的正则化参数调整方式, 使得正则化参数在迭代中稳定下降, 从而避免了在迭代中由于正则化参数调整不当而带来的多余正演计算过程。对于正演计算量很大的工程实际问题, 该方法在计算时间上的改进将会更加明显。算例 3 的计算结果表明, 在不同的噪声强度下, 本文提出的方法与梯度正则化法都能得到很好的计算结果,



即本文提出的方法也具有很好的抵抗观测噪声的能力。同时,从该算例还可看出,在不同的噪声强度下,采用本文方法的计算效率均明显优于梯度正则化法。

根据前面的计算结果可以看出,本文提出的大范围收敛梯度正则化算法收敛范围大,计算效率高,并具有一定的抵抗观测噪声的能力。此外,该方法实现简单,十分适合于不同领域非线性反问题的求解。

参考文献(References):

- [1] TIKHONOV A, ARSEN N V. *Solutions of Ill-Posed Problems* [M]. John Wiley & Sons, Inc 1977.
- [2] ENGL H W, HANKE M, NEUBAUER A. *Regularization of Inverse Problems* [M]. Kluwer, Dordrecht, 1996
- [3] HANKE M. A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems [J]. *Inverse Problems*, 1997, 3 (1): 79-95
- [4] KALTENBACHER B. Some Newton-type methods for the regularization of nonlinear inverse problems [J]. *Inverse Problems*, 1997, 13, (3): 729-753
- [5] BURGER M, MÜHLHUBER W. Iterative regularization of parameter identification problems by SQP methods [J]. *Inverse Problems*, 2002, 18 (4): 943-970
- [6] 唐立民, 张文飞, 刘迎曦. 微分方程反问题的梯度正则化方法 [J]. *计算结构力学及其应用*, 1991, 8 (2): 123-129 (TANG Limin, ZHANG Wen-fei, LIU Ying-xi Gradient-regularization method to solve the inverse problems of differential equations [J]. *Computational Structural Mechanics and Applications*, 1991, 8 (2): 123-129 (in Chinese))
- [7] 李晓江, 张文飞. 声波方程数值反演的 GR 方法 [J]. *石油地球物理勘探*, 1995, 30 (2): 201-206 (LI Xiaojiang, ZHANG Wen-fei Gradient regularization method for numerical inversion of acoustic wave equation [J]. *OGP*, 1995, 30 (2): 201-206 (in Chinese))
- [8] 刘迎曦, 王登刚, 张家良, 等. 材料物性参数识别的梯度正则化方法 [J]. *计算力学学报*, 2000, 17 (1): 69-75 (LIU Ying-xi, WANG Deng-gang, ZHANG Jialiang, et al Identification of material parameters with gradient-regularization method [J]. *Chinese Journal of Computational Mechanics*, 2000, 17 (1): 69-75 (in Chinese))
- [9] 王则柯, 高堂安. 同伦方法引论 [M]. 重庆: 重庆出版社, 1990 (WANG Ze-ke, GAO Tang-an. *Introduction of Homotopy Method* [M]. Chongqing: Chongqing Press, 1990 (in Chinese))
- [10] 王宇, 李兴斯. 解非线性极大极小问题的路径跟踪算法 [J]. *大连理工大学学报*, 1996, 36 (1): 117-119 (WANG Yu, LI Xing-si Path-following algorithm for solving nonlinear minimax problems [J]. *Journal of Dalian University of Technology*, 1996, 36 (1): 117-119 (in Chinese))

Global convergence gradient regularization algorithm for solving nonlinear inverse problems

CUI Kai¹, LI Xing-si¹, LI Bao-yuan², YANG Guo-wei¹

(1. Key Laboratory of High-Temperature Gas Dynamics, Institute of Mechanics, Chinese Academy of Science, Beijing 100080, China; 2. Department of Engineering Mechanics, Dalian University of Technology, Dalian 116023, China)

Abstract: Based on idea of homotopy mapping, an improved gradient regularization algorithm was developed. By using this path-following algorithm, the convergent bound of the gradient regularization method was efficiently widened. Moreover, a Sigmoid function was adopted to adjust the regularization parameter, by using this function, the efficiency and the stability of computation procedure were highly improved, while observational noises could also be resisted effectively. Numerical examples showed that the convergence bound of this algorithm is wider than normal gradient regularization algorithm, and the average efficiency is improved about 40-90%, besides, even though observational quantities were contaminated heavily by noise, an appropriate result could also be found.

Key words: inversion; gradient regularization method; homotopy method; regularization parameter