

地球物理反问题中两种正则化方法的比较

顾勇为¹ 归庆明¹ 边少锋² 郭建峰^{1,3}

(1 信息工程大学理学院数理系,郑州市 2201 信箱 160 号,450001)

(2 海军工程大学四系,武汉市解放大道 339 号,430033)

(3 中国科学院测量与地球物理研究所,武汉市徐东路 174 号,430077)

摘要:通过对最小二乘法的改造引入了两种正则化方法及其解的表达式,从理论上阐述了 Tikhonov 与 TS-VD 两种正则化方法效果的差异,特别就病态方程组在三种情况——系数阵良定秩、系数阵劣定秩右端项劣衰减、系数阵劣定秩右端项良衰减下正则化方法的应用效果作了比较,并给出了数值例子。

关键词:正则化方法;良定秩;劣定秩

中图法分类号:P223

由于地球物理本身的复杂性和不确定性,使地球物理反问题中存在许多不适定问题及病态方程,如位场延拓、波场反演、重磁解释、地震层析成像等大多归结为解病态的第一类算子方程,或者解其相应的病态离散代数方程组^[1,2,3]。为了解算这类病态方程,许多学者提出和发展了各种有效的数值方法,其中最著名的方法是 Tikhonov 正则化方法^[1,2,3]和截断奇异值分解(TSVD)正则化方法^[1,2,3]。

Tikhonov 正则化方法适用于解算各种病态方程,是一种非常通用且有效的方法,其关键是根据具体问题具体分析,然后选择合理的正则化矩阵和正则化参数^[4]。使用奇异值分解(SVD)法解算各种病态方程的关键是修改奇异值的技巧^[5],Wiggins^[6]采用了零化小奇异值的办法,即将小于某门限值的小奇异值化为 0,保留的非零奇异值个数即为数据的有效自由度,这就是著名的 TS-VD 正则化方法^[7]。

1 正则化方法的提出

众所周知,地球物理反问题往往归结为求解线性方程组的问题,即

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (1)$$

式中, \mathbf{A} 为 $m \times n$ 阶矩阵, $\text{rank}(\mathbf{A}) = n$ 。对地球物理反问题而言, \mathbf{A} 通常是接近奇异的。

直接使用广义逆反演方法求解式(1),得到未知参数 x 的最小二乘最小范数解为:

$$\mathbf{x}_{\text{LS}} = \mathbf{A}^+ \mathbf{b} \quad (2)$$

式中, \mathbf{A}^+ 是 \mathbf{A} 的 Moore-Penrose 广义逆。设 \mathbf{A} 的奇异值分解(SVD)式^[8]为:

$$\mathbf{A} = \mathbf{U}(\boldsymbol{\Sigma}, \mathbf{O})^T \mathbf{V}^T = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T \quad (3)$$

式中,正交矩阵 $\mathbf{U}_{m \times m} = (\mathbf{u}_1, \dots, \mathbf{u}_m)$; $\mathbf{V}_{n \times n} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$; $\boldsymbol{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_n)$, $\sigma_1 \geq \dots \geq \sigma_n > 0$ 是 \mathbf{A} 的奇异值。相应地, \mathbf{A}^+ 的奇异值分解式为:

$$\mathbf{A}^+ = \mathbf{V}(\boldsymbol{\Sigma}^{-1}, \mathbf{O})\mathbf{U}^T = \sum_{i=1}^n \sigma_i^{-1} \mathbf{v}_i \mathbf{u}_i^T \quad (4)$$

将式(4)代入式(2),得到最小二乘最小范数解的奇异值分解式为:

$$\begin{aligned} \mathbf{x}_{\text{LS}} &= \mathbf{A}^+ \mathbf{b} = \mathbf{V}(\boldsymbol{\Sigma}^{-1}, \mathbf{O})\mathbf{U}^T \mathbf{b} = \\ &= \sum_{i=1}^n \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i = \sum_{i=1}^n \frac{\boldsymbol{\alpha}_i}{\sigma_i} \mathbf{v}_i \end{aligned} \quad (5)$$

式中, $\boldsymbol{\alpha}_i = \mathbf{u}_i^T \mathbf{b}$ 。

当观测值 \mathbf{b} 表示为真值与噪声的和时,有 $\mathbf{b} = \mathbf{b}_t + \mathbf{e}$,可得式(5)的分解式为:

$$\mathbf{x}_{\text{LS}} = \sum_{i=1}^n \left(\frac{\mathbf{u}_i^T \mathbf{b}_t}{\sigma_i} + \frac{\mathbf{u}_i^T \mathbf{e}}{\sigma_i} \right) \mathbf{v}_i \quad (6)$$

分析式(6)可知,噪声对真解的污染主要反映在右端第二项,特别当奇异值 σ_i 异常小时,该污染项的值很大,甚至掩盖了反映真解的右端第一项。此时 \mathbf{A} 的条件数过大,为病态矩阵。

为克服 \mathbf{A} 的病态性, Tikhonov 正则化方法^[8,9]将式(5)或式(6)改造为:

$$\mathbf{x}_\lambda = \sum_{i=1}^n \frac{\alpha_i \sigma_i}{\lambda^2 + \sigma_i^2} \mathbf{v}_i = \sum_{i=1}^n \left(\frac{\sigma_i}{\lambda^2 + \sigma_i^2} \mathbf{u}_i^T \mathbf{b}_i + \frac{\sigma_i}{\lambda^2 + \sigma_i^2} \mathbf{u}_i^T \mathbf{e} \right) \mathbf{v}_i \quad (7)$$

其中, λ 为正则化参数。显然, 当 $\lambda = 0$ 时, $\mathbf{x}_\lambda = \mathbf{x}_{LS}$ 。适当选择正则化参数 λ , 可以减轻由于小 σ_i 放大噪声对解的污染, 但同时丧失了一些解的精确性。利用 TSVD 正则化方法^[8,9]将式(5)或式(6)改为:

$$\mathbf{x}_k = \sum_{i=1}^k \frac{\alpha_i}{\sigma_i} \mathbf{v}_i = \sum_{i=1}^k \left(\frac{\mathbf{u}_i^T \mathbf{b}_i}{\sigma_i} + \frac{\mathbf{u}_i^T \mathbf{e}}{\sigma_i} \right) \mathbf{v}_i \quad (8)$$

可以看出, 当 $k=n$ 时, $\mathbf{x}_k = \mathbf{x}_{LS}$ 。适当去除 $(n-k)$ 个大污染项, 恢复了一些解的主要特性, 同时也丧失了一些解的精确性。

2 正则化方法效果的比较

注意到式(1)中 \mathbf{x} 的最小二乘最小范数意义下的真解表达式(2), 将式(7)、式(8)作类似表示:

$$\mathbf{x}_\lambda = \mathbf{A}_\lambda^+ \mathbf{b} \quad (9)$$

$$\mathbf{x}_k = \mathbf{A}_k^+ \mathbf{b} \quad (10)$$

式中, $\mathbf{A}_\lambda^+ = \sum_{i=1}^n \frac{\sigma_i}{\lambda^2 + \sigma_i^2} \mathbf{v}_i \mathbf{u}_i^T$ 为 \mathbf{A} 的正则化广义逆^[8,9]; $\mathbf{A}_k^+ = \sum_{i=1}^k \sigma_i^{-1} \mathbf{v}_i \mathbf{u}_i^T$ 为 \mathbf{A}_k 的广义逆。

令 $\omega_k = \sigma_{k+1} / \sigma_k$, 则 ω_k^{-1} 为 TSVD 正则化方法中相邻两个保留与舍弃的奇异值之商, 称为奇异值的跨度(gap)^[7]。如果 \mathbf{A} 的奇异值 $\sigma_1, \sigma_2, \dots, \sigma_n$ 可明显分为两组: $\sigma_1, \sigma_2, \dots, \sigma_k$ 和 $\sigma_{k+1}, \sigma_{k+2}, \dots, \sigma_n$, 同一组奇异值之间在数量级上基本无差异, 而两组奇异值之间在数量级上差异很大, 则称 \mathbf{A} 为良定秩, 否则, 称 \mathbf{A} 为劣定秩。可以看出, \mathbf{A} 为良定秩时, 奇异值跨度 ω_k^{-1} 很大, \mathbf{A} 为劣定秩时很小。

定理 1^[7] 若取 $\lambda \in [\sigma_{k+1}, \sigma_k]$, 则有:

$$\|\mathbf{A}_k^+\|/2 \leq \|\mathbf{A}_\lambda^+\| \leq \|\mathbf{A}_k^+\|/2\omega_k \quad (11)$$

特别地, 当 $\lambda = \sigma_k$ 时, $\|\mathbf{A}_\lambda^+\| = \|\mathbf{A}_k^+\|/2$; 当 $\lambda = \sigma_{k+1}$ 时, $\|\mathbf{A}_\lambda^+\| = \|\mathbf{A}_k^+\|/2\omega_k$ 。

该定理表明, 在 \mathbf{A} 为良定秩的情况下, ω_k^{-1} 很大, 若选取 λ 靠近 σ_{k+1} , 则两种正则化方法的差异会很大; 但若选取 λ 靠近 σ_k , 则两者很接近。这就为选择正则化参数 λ 提供了一个参照系。

定理 2^[7] 若取 $\lambda \in [\sigma_{k+1}, \sigma_k]$, 则有:

$$\omega_k^{1/2} / (1 + \omega_k^{1/2}) \leq$$

$$\min_{\lambda} \|\mathbf{A}_\lambda^+ - \mathbf{A}_k^+\| / \|\mathbf{A}_k^+\| \leq \omega_k^{1/2} / (1 + \omega_k^{3/2}) \quad (12)$$

该定理表明, 在 \mathbf{A} 为良定秩的情况下, 由于 $\omega_k \approx 0$, \mathbf{A}_λ^+ 与 \mathbf{A}_k^+ 的相对误差较小, \mathbf{x}_λ 与 \mathbf{x}_k 是接近的。还可以证明, 当 $\lambda = (\sigma_k \sigma_{k+1})^{1/2}$ 时, $\|\mathbf{A}(\mathbf{A}_\lambda^+ - \mathbf{A}_k^+)\|$ 可达到最小。 $\min_{\lambda} \|\mathbf{A}(\mathbf{A}_\lambda^+ - \mathbf{A}_k^+)\|$ 的意义在于两种正则化方法的残差之差此时达到最小:

$$(\mathbf{b} - \mathbf{A}\mathbf{x}_k) - (\mathbf{b} - \mathbf{A}\mathbf{x}_\lambda) = \mathbf{A}(\mathbf{A}_\lambda^+ - \mathbf{A}_k^+)\mathbf{b}$$

3 劣定秩下两种正则化方法的比较

劣定秩下两种正则化方法分别为:

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}_\lambda\|^2 = \sum_{i=1}^n (\lambda^2 / (\sigma_i^2 + \lambda^2))^2 (\mathbf{u}_i^T \mathbf{b})^2 \quad (13)$$

$$\|\mathbf{b} - \mathbf{A}\mathbf{x}_k\|^2 = \sum_{i=k+1}^n (\mathbf{u}_i^T \mathbf{b})^2 \quad (14)$$

由此可见, 如果 $\mathbf{u}_i^T \mathbf{b}$ 衰减得足够快, 则两种正则解对应的残差都很小。

定理 3 如果有 $\mathbf{u}_i^T \mathbf{b} = \sigma_i^2$, $\omega_k \ll 1$, $\lambda \in [\sigma_{k+1}, \sigma_k]$, 那么, 有:

$$\sigma_k \frac{\omega_k^{3/2}}{1 + \omega_k^{1/2}} \leq \min_{\lambda} \|x_\lambda - x_k\|_\infty \leq \sigma_k \frac{\omega_k^{3/2}}{1 + \omega_k^{3/2}} \quad (15)$$

证明:

$$\min_{\lambda} \|\mathbf{x}_\lambda - \mathbf{x}_k\|_\infty =$$

$$\min_{\lambda} \left\| \sum_{i=1}^n \frac{\mathbf{u}_i^T \mathbf{b} \sigma_i}{\sigma_i^2 + \lambda^2} \mathbf{v}_i - \sum_{i=1}^k \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i \right\|_\infty = \min_{\lambda} \left\{ \max \left(\frac{\sigma_k \lambda^2}{\sigma_k^2 + \lambda^2}, \frac{\sigma_{k+1}^3}{\sigma_{k+1}^2 + \lambda^2} \right) \right\}$$

式中, $\sigma_i \lambda^2 / (\sigma_i^2 + \lambda^2)$ 为 λ^2 的增函数; $\sigma_i^3 / (\sigma_i^2 + \lambda^2)$ 为 λ^2 的减函数。通过分析可以看出, $\min_{\lambda} \left\{ \max \left(\frac{\sigma_k \lambda^2}{\sigma_k^2 + \lambda^2}, \frac{\sigma_{k+1}^3}{\sigma_{k+1}^2 + \lambda^2} \right) \right\}$ 在 λ^2 恰好位于两函数的交点处达到, 由此可解得:

$$\lambda^2 = \frac{\sigma_{k+1}^3 - \sigma_{k+1}^2 \sigma_k + \sqrt{\sigma_{k+1}^4 (\sigma_k - \sigma_{k+1})^2 + 4 \sigma_{k+1}^3 \sigma_k^3}}{2 \sigma_k}$$

因此,

$$\sigma_k \frac{\omega_k^{3/2}}{1 + \omega_k^{1/2}} \leq \min_{\lambda} \|\mathbf{x}_\lambda - \mathbf{x}_k\|_\infty = \sigma_k \frac{1}{1 + \frac{1}{\lambda^2 / \sigma_k^2}} \leq \sigma_k \frac{\omega_k^{3/2}}{1 + \omega_k^{3/2}}$$

该定理表明, 即使在 \mathbf{A} 为劣定秩的情况下, 只要 $\mathbf{u}_i^T \mathbf{b}$ 衰减足够快(按奇异值平方衰减), 并且能找到某个 k , 使得 $\omega_k \ll 1$, 就能保证由两种正则化方法获得的解近似相等。下述定理更为广泛地说明了这个事实, 证法与定理 3 类似。

定理 4 如果有 $u_i^T b = \sigma_i^m (m > 1), \omega_k \ll 1, \lambda \in [\sigma_{k+1}, \sigma_k]$, 那么, 有:

$$\min_{\lambda} \|x_{\lambda} - x_k\|_{\infty} = \begin{cases} \sigma_k O(\omega_k^{(m+1)/2}), & 1 < m < 3 \\ \sigma_k O(\omega_k^2), & m \geq 3 \end{cases} \quad (16)$$

4 算例分析

4.1 系数阵劣定秩且右端项劣衰减情况

笔者采用 11×9 的 Hilbert 矩阵作为方程组 (1) 的系数矩阵, 该矩阵的最大奇异值为 1.7, 最小奇异值为 1.7×10^{-11} , 条件数为 10^{11} , 可见其病态程度相当严重, 分析 9 个奇异值, 没有明显分为跨度较大的两组, 可知 A 是劣定秩的。先给出方程组的准确解, 再由式 (1) 算出右端项 b , 然后将 b 作 $10^{-3} \times \text{rand}(11, 1)$ 的扰动, 于是得到了一个具有劣定秩系数阵且右端项劣衰减的病态方程组。分别用 SVD 法、TSVD 正则化方法 (取 $k=5$) 和 Tikhonov 正则化方法 (取 $\lambda = (\sigma_5 \sigma_6)^{1/2}$) 解该方程组, 其结果如表 1 所示。

表 2 第二种情况下三种解法的结果比较

Tab. 2 Comparison Between Three Solutions in Case 2

x _i 序号	准确解	SVD 解	TSVD 解	Tikhonov 解
1	-1.001 7	9.767 2×10 ¹	-1.001 8	-1.005 5
2	-0.798 6	-6.013 5×10 ³	-0.798 9	-0.793 0
3	-0.658 0	9.748 0×10 ⁴	-0.656 8	-0.652 9
4	-0.559 0	-6.688 3×10 ⁵	-0.557 8	-0.556 1
5	-0.485 8	2.372 4×10 ⁶	-0.485 1	-0.485 1
6	-0.429 5	-4.707 5×10 ⁶	-0.429 5	-0.430 7
7	-0.384 9	5.274 5×10 ⁶	-0.385 5	-0.387 6
8	-0.348 6	-3.118 1×10 ⁶	-0.349 8	-0.352 5
9	-0.318 5	7.559 3×10 ⁵	-0.320 3	-0.323 3
最大误差	0.000 0	5.274 5×10 ⁶	0.001 2	0.005 5
平均误差	0.000 0	1.889 0×10 ⁶	8.185 1×10 ⁻⁴	0.003 4
均方误差	0.000 0	7.373 3×10 ¹²	9.761 5×10 ⁻⁶	1.407 7×10 ⁻⁵

4.3 系数阵良定秩情况

将 Hilbert 矩阵的 9 个奇异值略作修改, 改变第 4 个和第 5 个奇异值, 使前 4 个奇异值和后 5 个奇异值出现明显的跨度。仍先给准确解, 算出 b 并作扰动, 得到一个具有良定秩系数阵的病态方程组。分别用 SVD 法、TSVD 正则化方法 (取 $k=4$) 和 Tikhonov 正则化方法 (取 $\lambda = (\sigma_4 \sigma_5)^{1/2}$) 解该方程组, 结果如表 3 所示。

5 结 语

理论分析和数值计算取得了一致的结论。在系数阵劣定秩且右端项劣衰减情况下, 方程组的

表 1 第一种情况下三种解法结果比较

Tab. 1 Comparison Between Three Solutions in Case 1

x _i 序号	准确解	SVD 解	TSVD 解	Tikhonov 解
1	1.0	-4.684 0×10 ¹	1.022 5	0.987 4
2	1.0	2.937 9×10 ³	0.720 7	1.069 8
3	1.0	-4.532 0×10 ⁴	1.575 2	0.963 4
4	1.0	3.000 3×10 ⁵	1.093 0	0.947 4
5	1.0	-1.033 3×10 ⁶	0.695 4	0.966 6
6	1.0	1.999 8×10 ⁶	0.610 7	0.992 6
7	1.0	-2.192 9×10 ⁶	0.771 4	1.015 0
8	1.0	1.272 2×10 ⁶	1.078 8	1.030 6
9	1.0	-3.034 3×10 ⁵	1.457 7	1.039 3
最大误差	0.0	2.192 9×10 ⁶	0.575 2	0.069 8
平均误差	0.0	7.944 2×10 ⁵	0.269 9	0.033 0
均方误差	0.0	1.297 6×10 ¹²	0.103 4	0.001 4

4.2 系数阵劣定秩且右端项良衰减情况

对方程组 (1) 改变右端项 b , 使其满足右端项良衰减条件 ($u_i^T b = \sigma_i^2, \omega_k \ll 1$), 此时真解为 $x = V \cdot (\sigma_1, \sigma_2, \dots, \sigma_9)^T$, 由此真解及 A 算出 b 后作扰动 (方法同上), 于是得到一个系数阵劣定秩且右端项良衰减的病态方程组。分别用 SVD 法、TSVD 正则化方法 (取 $k=4$) 和 Tikhonov 正则化方法 (取 $\lambda = (\sigma_4 \sigma_5)^{1/2}$) 解该方程组, 其结果如表 2 所示。

表 3 第三种情况下三种解法结果比较

Tab. 3 Comparison Between Three Solutions in Case 3

x _i 序号	准确解	SVD 解	TSVD 解	Tikhonov 解
1	1.0	-9.560 7×10 ¹	1.000 9	0.991 6
2	1.0	6.214 5×10 ³	1.001 2	1.006 1
3	1.0	-9.944 5×10 ⁴	0.987 5	1.016 5
4	1.0	6.783 5×10 ⁵	0.998 0	1.023 1
5	1.0	-2.396 0×10 ⁶	1.008 5	1.019 8
6	1.0	4.738 9×10 ⁶	1.012 0	1.008 1
7	1.0	-5.296 1×10 ⁶	1.008 4	0.990 3
8	1.0	3.124 3×10 ⁶	0.998 8	0.968 6
9	1.0	-7.561 9×10 ⁵	0.984 9	0.944 7
最大误差	0.0	5.296 1×10 ⁶	0.015 1	0.023 1
平均误差	0.0	1.899 5×10 ⁶	0.006 9	0.019 8
均方误差	0.0	7.459 9×10 ¹²	7.534 8×10 ⁻⁵	0.000 6

TSVD 解和 Tikhonov 解都比 SVD 解准确,但都与真解相差较大;在系数阵劣定秩且右端项良衰减情况下,方程组的 TSVD 解和 Tikhonov 解都比 SVD 解更加准确,且都与真解具有较小的误差;在系数阵良定秩的情况下,方程组的 TSVD 解和 Tikhonov 解十分接近,且都与真解具有较小的误差。因此,在后两种情况下,用 TSVD 正则化方法和 Tikhonov 正则化方法是可以信赖的。

参 考 文 献

1 栾文贵. 地球物理中的反问题. 北京:科学出版社, 1989

2 杨文采. 地球物理反演和地震层析成像. 北京:地质出版社,1989

3 傅淑芳, 朱仁益. 地球物理反问题. 北京:地震出版社,1998

4 唐隆基,李 文,邓阳生. 关于地球物理中病态方程的

若干问题. 地球物理学报,1995,38(1):105~114

5 李 平,王椿镛,许厚泽,等. 地球物理反演中奇异值分解应用的若干问题探讨. 自然科学进展,2001,11(8):891~896

6 Wiggins R A. The General Inverse Problem: Implication of Surface Waves and Free Oscillation for Earth Structure. Rev. Geophys. Space Phys., 1972, 10: 251~285

7 Hansen P C. The Truncated SVD as a Method for Regularization. BIT,1987, 27: 534~553

8 Tikhonov A N, Arsenin V Y. Solutions of Ill-Posed Problems. New York: Wiley, 1977

9 Hansen P C. Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion. Philadelphia, SIAM, 1998

第一作者简介:顾勇为,副教授。现主要从事应用数学与物理大地测量等研究。
E-mail: gyw1019@sina.com

Comparison Between Tikhonov Regularization and Truncated SVD in Geophysics

GU Yongwei¹ GUI Qingming¹ BIAN Shaofeng² GUO Jianfeng^{1,3}

(1 Department of Mathematics and Physics, Information Engineering University, 160 Mail-box 2201, Zhengzhou 450001, China)

(2 Institute of Navigation, Naval University of Engineering, 339 Jiefang Avenue, Wuhan 430033, China)

(3 Institute of Geodesy and Geophysics, Chinese Academy of Sciences, 174 Xudong Road, Wuhan 430077, China)

Abstract: Ill-posed problem often appears in the inverse problem in geophysics. Many regularization methods are used to produce reasonable solutions. Among them, Tikhonov regularization and truncated SVD (Singular Value Decomposition) are two important methods used widely, but how to compare them is a challenging and meaningful area under discussion. The paper discusses the comparison between them. First, with the SVD of coefficients matrix, the ordinary least squares (LS) solution to linear equation is expressed by decomposition of pseudo-inverse of the matrix. Then, through the analysis of the parameters, the similarity and difference between the two methods are declared theoretically. Last, some numerical examples are given to confirm the theory.

Key words: regularization; well-defined rank; bad determined rank

About the first author: GU Yongwei, associate professor. His major research interest includes application of mathematics and survey in geophysics, etc.
E-mail: gyw1019@sina.com