

摘 要

为处理噪声数据，最近吉洪诺夫正则化的近似多项式及其误差估计已被推导出来。基于已经得出的 L_2 误差界和一致误差界，本文首先对这两种误差界中的某些项进行量化，给出量化后的 L_2 误差界和一致误差界。本文通过分析这两种误差估计，可知吉洪诺夫正则化的近似多项式中所含的正则化参数能降低噪声，因此本文还对此公式中的正则化参数进行了选取。本文采用了两种方法选取正则化参数，分别是平衡原则和 Brezinski–Rodriguez–Seatzu 估计法。数值实验表明两种方法选取的正则化参数是有效的，即在一定程度上能降低噪声。最后，本文还分析了吉洪诺夫正则化的改进的 Lagrange 插值公式和吉洪诺夫正则化的重心插值公式的数值稳定性，得出这两个吉洪诺夫正则化的插值公式都是向前稳定的，其中吉洪诺夫正则化的改进的 Lagrange 插值公式还是向后稳定的。数值实验也验证了这两个插值公式的数值稳定性结论。

关键词：吉洪诺夫正则化，正则化参数选取，数值稳定性，平衡原则，Brezinski–Rodriguez–Seatzu 估计法

Abstract

To deal with noisy data, recently Tikhonov regularized approximation polynomial and its error estimates have been derived. Based on the already derived L_2 error bound and the uniform error bound, this thesis first quantifies certain terms in these two error bounds to give the quantified L_2 error bound and the uniform error bound. By these two error estimates it is also explained that the regularization parameter contained in the Tikhonov regularized approximation polynomial can reduce noise. Therefore, the regularization parameter in this formula are also selected in this thesis. Two methods are used to select the regularization parameter, which are the balancing principle and the Brezinski–Rodriguez–Seatzu estimators. Numerical experiments show that the regularization parameter selected by the two methods is effective, that is, it can reduce noise to some extent. Finally, the numerical stability of the Tikhonov regularized modified Lagrange interpolation formula and the Tikhonov regularized barycentric interpolation formula are investigated. The analysis in this thesis concludes that both Tikhonov regularized interpolation formulae are forward stable, of which the Tikhonov regularized modified Lagrange interpolation formula is still backward stable. The numerical experiments also verify the theoretical results of the numerical stability of these two interpolation formulae.

Keywords: Tikhonov regularization, parameter choice strategy, numerical stability, balancing principle, Brezinski–Rodriguez–Seatzu estimators

目 录

摘要	I
Abstract	II
1. 绪论	1
1.1 研究背景	1
1.2 研究现状	4
1.2.1 正则化参数选取的研究现状	4
1.2.2 数值稳定性的研究现状	5
1.3 本文的创新点	6
1.4 本文的结构安排	6
2. 吉洪诺夫正则化的近似多项式的误差估计	8
2.1 预备知识	8
2.1.1 L_2 误差	8
2.1.2 一致误差	9
2.2 误差估计的两个推论	10
2.2.1 L_2 误差的推论	11
2.2.2 一致误差的推论	12
2.3 本章小结	13
3. 吉洪诺夫正则化的近似多项式的参数选取	14
3.1 平衡原则	14
3.2 Brezinski-Rodriguez-Seatzu 估计法	17

3.2.1	预备知识	17
3.2.2	BRS 估计法的应用	18
3.3	数值实验	19
3.4	本章小结	22
4.	吉洪诺夫正则化的插值公式的数值稳定性	24
4.1	预备知识	24
4.2	吉洪诺夫正则化的插值公式的数值稳定性	26
4.2.1	吉洪诺夫正则化的改进的 Lagrange 插值公式的数值稳定性	27
4.2.2	吉洪诺夫正则化的重心插值公式的数值稳定性	27
4.3	数值实验	29
4.4	本章小结	31
5.	总结与展望	32
5.1	全文总结	32
5.2	后续工作展望	32
	参考文献	34
	致谢	40
	攻读硕士学位期间的研究成果	41

1. 绪论

1.1 研究背景

给定一个较复杂的函数 f , 为了方便计算和研究, 通常会用一个较简单的函数对 f 进行逼近, 而多项式是用于逼近函数的常见形式。多项式逼近应用广泛, 在插值、数值积分、微分方程的数值解和积分方程的数值解等领域中, 多项式逼近都会作为一个基本手段, 尤其是以正交多项式为基的级数展式的逼近。下面给出文献 [1] 中以正交多项式为基的多项式逼近。

设 $\{\Phi_\ell(x)\}_{\ell=0}^\infty$ 是正交多项式族, 其中 ℓ 是正交多项式 $\Phi_\ell(x)$ 的次数。设在 $[-1, 1]$ 上的连续函数空间为 $\mathcal{C}([-1, 1])$, 对函数 $f(x) \in \mathcal{C}([-1, 1])$ 进行逼近的 L 次多项式为

$$p_L(x) = \sum_{\ell=0}^L \beta_\ell \tilde{\Phi}_\ell(x), \quad x \in [-1, 1], \quad (1.1)$$

其中 $\{\beta_\ell\}_{\ell=0}^L$ 是待确定的系数集, $\tilde{\Phi}_\ell(x)$ 是 ℓ 次的标准正交多项式, 即

$$\tilde{\Phi}_\ell(x) := \frac{\Phi_\ell(x)}{\|\Phi_\ell(x)\|_{L_2}}, \quad \ell = 0, \dots, L,$$

其中 L_2 范数

$$\|f\|_{L_2} := \sqrt{\langle f(x), f(x) \rangle_{L_2}} = \left(\int_{-1}^1 w(x) |f(x)|^2 dx \right)^{1/2}, \quad (1.2)$$

其中 $w(x)$ 是非负权函数, 满足

$$\int_{-1}^1 w(x) dx < \infty,$$

其中 L_2 内积为 $\langle f(x), g(x) \rangle_{L_2} := \int_{-1}^1 w(x) f(x) g(x) dx$.

为了确定系数 $\{\beta_\ell\}_{\ell=0}^L$, 如果考虑用离散的方式研究这里的多项式逼近问题, 那么需要样本数据。然而在实践中, 样本数据通常会被噪声污染, 并且考虑到经典的最小二乘法对噪声数据较敏感, 因此可以通过正则化来降低噪

声。吉洪诺夫正则化^[2, 3]是一种常用的正则化方法，它在反问题和不适定问题中得到了广泛的研究^[4-9]。基于上述情况，下面给出用吉洪诺夫正则化确定系数 $\{\beta_\ell\}_{\ell=0}^L$ 的结果，从而得出近似多项式 $p_L(x)$ ，详细过程参考文献 [1]。

假设有 $N + 1$ 个样本数据，设 $\mathcal{X}_{N+1} = \{x_0, x_1, \dots, x_N\}$ 是在 $[-1, 1]$ 上的 $N + 1$ 次正交多项式 $\Phi_{N+1}(x)$ 的零点集 ($N + 1$ 次的正交多项式的零点就是 $N + 1$ 个高斯求积点^[10, 11])，对应的高斯求积系数为 $\{w_0, w_1, \dots, w_N\}$ 。函数 $f(x) \in C([-1, 1])$ 在 \mathcal{X}_{N+1} 上的值为 $\{f(x_j)\}$ (可能含噪声)，则上述问题对应的吉洪诺夫正则化问题为

$$\min_{\beta_\ell \in \mathbb{R}} \left\{ \sum_{j=0}^N w_j \left(\sum_{\ell=0}^L \beta_\ell \tilde{\Phi}_\ell(x_j) - f(x_j) \right)^2 + \lambda \sum_{\ell=0}^L |\beta_\ell|^2 \right\}, \quad (1.3)$$

其中 \mathbb{R} 是实数集， $\lambda > 0$ 是正则化参数。文献 [1] 得出，若 L 和 N 满足 $2L \leq 2N + 1$ ，则吉洪诺夫正则化问题(1.3)的最优解为

$$\beta_\ell = \frac{1}{1 + \lambda} \sum_{j=0}^N w_j \tilde{\Phi}_\ell(x_j) f(x_j), \quad \ell = 0, 1, \dots, L, \quad \lambda > 0. \quad (1.4)$$

将 β_ℓ 的表达式(1.4)代入到(1.1)式中，可以得到吉洪诺夫正则化的近似多项式

$$p_{L, N+1}(x) = \frac{1}{1 + \lambda} \sum_{\ell=0}^L \left(\sum_{j=0}^N w_j \tilde{\Phi}_\ell(x_j) f(x_j) \right) \tilde{\Phi}_\ell(x), \quad \lambda > 0. \quad (1.5)$$

文献 [1] 还对吉洪诺夫正则化的近似多项式(1.5)的逼近质量进行了研究，并且分析得出(1.5)式中的正则化参数 λ 的取值非常关键。如果正则化参数 λ 取值太大，那么多项式逼近的效果不好，如果 λ 取值太小，相当于没有进行正则化处理，达不到降低噪声的目的。在文献 [1] 的数值实验部分，正则化参数 λ 是通过逐一试验 $\lambda = 10^{-2}, 10^{-1.9}, \dots, 10^{-0.1}, 1$ 来选取的最佳参数值。这种方式虽然可行，但是低效。因为需要试验多次，并且需要预估最佳 λ 的取值范围。因此，本文研究的内容之一是如何选取(1.5)式中的正则化参数 λ 。

另外，当 $L = N$ 时，即在 $N + 1$ 个点上得出 N 次的吉洪诺夫正则化的近似多项式(1.5)还可以被改写成两种吉洪诺夫正则化的插值公式^[1]。这与两种经典的重心插值公式有关，分别是改进的 Lagrange 插值公式

$$p_N^{\text{mdf}}(x) = \ell(x) \sum_{j=0}^N \frac{\Omega_j}{x - x_j} f(x_j), \quad (1.6)$$

和重心插值公式

$$p_N^{\text{bary}}(x) = \frac{\sum_{j=0}^N \frac{\Omega_j}{x-x_j} f(x_j)}{\sum_{j=0}^N \frac{\Omega_j}{x-x_j}}, \quad (1.7)$$

其中

$$\ell(x) = (x - x_0)(x - x_1) \cdots (x - x_N), \quad (1.8)$$

重心权

$$\Omega_j = \frac{1}{\prod_{k \neq j} (x_j - x_k)}, \quad j = 0, 1, \dots, N. \quad (1.9)$$

两个插值公式(1.6)和(1.7)的推导细节见文献 [12].

对于吉洪诺夫正则化的近似多项式(1.5), 当 $L = N$ 时, 设插值节点是 $N + 1$ 个高斯求积点 $\{x_j\}_{j=0}^N$, 对应的函数值为 $\{f(x_j)\}_{j=0}^N$. 通过 Christoffel-Darboux 公式^[13] 以及重心权与高斯求积系数的关系式^[14]

$$\Omega_j = w_j \tilde{\Phi}_N(x_j),$$

吉洪诺夫正则化的近似多项式(1.5)可以转化成两个吉洪诺夫正则化的插值公式, 分别为吉洪诺夫正则化的改进的 Lagrange 插值公式^[1]

$$p_N^{\text{Tik-mdf}}(x) = \frac{\ell(x)}{1 + \lambda} \sum_{j=0}^N \frac{\Omega_j}{x - x_j} f(x_j), \quad \lambda > 0, \quad (1.10)$$

其中 $\ell(x)$ 的定义见(1.8)式, 权 Ω_j 的定义见(1.9)式, 以及吉洪诺夫正则化的重心插值公式^[1]

$$p_N^{\text{Tik-bary}}(x) = \frac{\sum_{j=0}^N \frac{\Omega_j}{x-x_j} f(x_j)}{(1 + \lambda) \sum_{j=0}^N \frac{\Omega_j}{x-x_j}}, \quad \lambda > 0, \quad (1.11)$$

其中权 Ω_j 的定义见(1.9)式。

由此可以看出, 吉洪诺夫正则化的两个插值公式(1.10)和(1.11)都是在经典版的插值公式(1.6)和(1.7)上增加了一个乘数因子 $1/(1 + \lambda)$. 如果令 $\lambda = 0$, 那么吉洪诺夫正则化的两个插值公式(1.10)式和(1.11)式分别转化为经典版的改进的 Lagrange 插值公式(1.6)和重心插值公式(1.7). 与吉洪诺夫正则化的近似多项式(1.5)中的 λ 一样, 吉洪诺夫正则化的两个插值公式(1.10)式和(1.11)式中的正则化参数 λ 可以降低数据中的噪声, 但是经典版的插值公式(1.6)和(1.7)却不能, 见文献 [1] 的数值实验部分. 因此运用吉洪诺夫正则化

的插值公式计算是重要的，吉洪诺夫正则化的插值公式的运算结果精确性与它的数值稳定性密不可分。在算法执行过程中，若数值稳定性好，即舍入误差对计算结果的精确性影响较小；反之，舍入误差对计算结果的精确性影响较大。因此研究吉洪诺夫正则化的两个插值公式(1.10)和(1.11)的数值稳定性是有必要的，这也是本文关注的另一个问题。由于吉洪诺夫正则化的两个插值公式(1.10)和(1.11)与经典版的两个插值公式(1.6)和(1.7)在形式上只相差一个乘数因子 $1/(1 + \lambda)$ ，因此吉洪诺夫正则化版的插值公式的数值稳定性分析可以结合 Higham^[15] 对(1.6)式和(1.7)式的数值稳定性的研究。

1.2 研究现状

1.2.1 正则化参数选取的研究现状

正则化参数选取的问题一直以来都是正则化研究中的一个重要课题，目前存在许多种选取正则化参数的方法。这些方法主要分为两大类，先验选取和后验选取。结合吉洪诺夫正则化的先验选取方法已经有许多结果，在某些条件下，收敛速度的结果也得到了证明^[16-19]。先验选取往往具有理论分析价值，然而在实际应用中如何选取正则化参数通常是不清楚的，因为通常难以检验它选取的条件。因此后验选取方法显得尤为重要，下面介绍一些常见且实用的后验选取方法。

在噪声水平未知的情况下，被广泛使用的后验选取策略有 L 曲线法^[20] 和广义交叉验证法 (GCV)^[21]。L 曲线法是在横纵坐标取对数的尺度下，绘制残差范数和正则化解范数在一组正则化参数下的图像，得到的曲线形状通常像字母 L，所以 Hansen 将此方法称为 L 曲线法。L 曲线法的使用就是找到该 L 曲线上的“角点”所对应的正则化参数，此参数值就是选取的最优正则化参数，此“角点”也就是该 L 曲线中曲率最大的点。但是采用 L 曲线法并不总能找到合适的正则化参数，因为得到的曲线可能出现多个“角点”或者一个“角点”也没有的情况^[22]。此外，L 曲线法是不收敛的^[23]，并且对于解决大型问题而言，计算量较大。GCV 方法是给出一个关于正则化参数 λ 的 GCV 函数 $G(\lambda)$ ，找到使 $G(\lambda)$ 取最小值时所对应的正则化参数 λ 的值，就是此方法确定的最优正则化参数。根据 Varah^[24] 的研究表明，当使用 GCV 方法时，GCV 函数通常难以求得最小值^[22]，因为函数在最小值附近相当平坦，从而导致 GCV 方法选的正则化参数不准确。随后，Brezinski, Rodriguez 和 Seatzu 三位学者

提出了另一种正则化参数选取的方法, 即 Brezinski-Rodríguez-Seatzu(简称 BRS) 估计法, 该方法也不需要提前知道数据的噪声水平。在文献 [25] 的数值实验部分将 BRS 估计法与 L 曲线法和 GCV 方法进行比较, 结果 BRS 估计法表现更好。本文选取(1.5)式中的正则化参数 λ 用的方法之一是 BRS 估计法。另外, 不需要已知噪声水平的方法还有 Hanke-Raus 规则^[26], 拟最优性准则^[27] 等。

在噪声水平已知的情况下, 平衡原则^[3, 28-30] 是一种有效且实用的后验选取方法。这一方法起源于 Lepskii^[31] 关于非参数回归估计的研究。随后, 在反问题的背景下, 这一方法在文献 [32] 中被引入。此后, 许多学者对各种正则化技术采用了这一方法来选取正则化参数, 例如 Banach 空间中梯形法则的自正则化和非线性问题的 Lavrentiev 正则化^[28]。Pereverzev 等^[33] 将此方法用于球面上的多项式中的参数选取, 这与本文用平衡原则对多项式(1.5)中参数选取的方法类似。另外, 需要已知噪声水平的后验选取方法还有 Morozov 偏差原理^[34-36], 这也是一种十分有效的选取正则化参数的方法。Bonesky^[37] 将 Morozov 偏差原理用于吉洪诺夫正则化的参数选取, 并研究了正则化性质以及收敛速度。Lu 等^[38, 39] 将这一方法用于多正则化参数的选取, 并给出了相应的算法。需要已知噪声水平的后验选取方法还有 Engle 准则^[40], 单调误差规则^[41], Arcangeli 准则^[42] 等。

1.2.2 数值稳定性的研究现状

插值公式的数值稳定性一直备受关注。Rack 和 Reimer^[43] 首先对改进的 Lagrange 插值公式(1.6)进行数值稳定性研究, 他们给出了舍入误差分析, 得出了一个较弱的边界, 但没有确定改进的 Lagrange 插值公式(1.6)的向后稳定性。为此, 在文献 [15] 中, Higham 进一步研究了改进的 Lagrange 插值公式(1.6)和重心插值公式(1.7)的数值稳定性, 得出改进的 Lagrange 插值公式(1.6)在任何插值节点都是向后稳定的 (包括等距节点), 而重心插值公式(1.7)仅在 Lebesgue 常数小的插值节点上是向前稳定的。若更形象地描述这类插值节点, 就是插值节点在区间两端分布比较密集, 在中间分布比较稀疏。满足条件的插值节点有第一类 Chebyshev 点 (Chebyshev 多项式的零点) 和第二类 Chebyshev 点 (Chebyshev 多项式的极点)。随后, 对于改进的 Lagrange 插值公式和重心插值公式, Webb 和 Trefethen^[44] 研究了它们外推时的数值稳定性。研究表明, 对于在复平面上的点的计算, 重心插值公式变得

不稳定, 应该被改进的 Lagrange 插值公式取代。此外, Mascarenhas^[45] 对在第二类 Chebyshev 点上插值的改进的 Lagrange 插值公式和重心插值公式的数值稳定性作了进一步的分析, 并表明在第二类 Chebyshev 点上的重心插值公式比之前认为的更加稳定。

1.3 本文的创新点

本文的创新点主要有以下几个方面:

(1) 本文对吉洪诺夫正则化的近似多项式(1.5)的 L_2 误差界和一致误差界的某些项进行量化, 给出量化后的 L_2 误差和一致误差的推论。

(2) 本文对吉洪诺夫正则化的近似多项式(1.5)中的正则化参数 λ 进行选取。由 L_2 误差估计和一致误差估计可以分析得出正则化参数 λ 的取值很关键, 它影响着逼近的效果和数据的降噪。本文将采用两种选取正则化参数的方法, 分别是平衡原则和 Brezinski-Rodriguez-Seatzu 估计法。数值实验表明这两种方法可以选出合适的正则化参数, 从而达到降噪的目的。另外, 通过对振荡函数的测试, 本文还比较了这两种选取正则化参数的方法。

(3) 本文分析了吉洪诺夫正则化的改进的 Lagrange 插值公式(1.10)和吉洪诺夫正则化的重心插值公式(1.11)的数值稳定性。结合 Higham 对改进的 Lagrange 插值公式和重心插值公式的稳定性研究, 得出了吉洪诺夫正则化的两个插值公式的数值稳定性结论。最后用数值实验进行验证, 数值稳定性的实验结果与理论结果一致。

1.4 本文的结构安排

本文分为五个部分。第一部分是绪论, 介绍了研究背景以及正则化参数选取的研究现状和数值稳定性的研究现状。第二部分对吉洪诺夫正则化的近似多项式的 L_2 误差估计和一致误差估计中的某些项进行量化, 得出了两个推论。第三部分对(1.5)式中的正则化参数 λ 进行选取, 使用的是平衡原则和 BRS 这两种选取正则化参数的方法。数值实验表明, 这两种正则化参数选取的方法是有效的, 且本文对这两种方法也作了数值实验比较。第四部分研究了吉洪诺夫正则化的改进的 Lagrange 插值公式(1.10)和吉洪诺夫正则化的重心插值公式(1.11)的数值稳定性。理论分析得出这两个吉洪诺夫正则化的插值公式都是向前稳定的, 并且吉洪诺夫正则化的改进的 Lagrange 插值公

式(1.10)也是向后稳定的。数值结果也验证了这一数值稳定性的结论。第五部分是对本文工作的总结以及对后续工作的展望。

2. 吉洪诺夫正则化的近似多项式的误差估计

2.1 预备知识

这一部分主要来自文献 [1]. 在噪声存在的情况下, 吉洪诺夫正则化的近似多项式(1.5)有两种误差估计, 分别是 L_2 误差估计和一致误差估计. 在给出这两种误差估计之前, 首先作一些符号说明. 设函数 $f \in \mathcal{C}([-1, 1])$, f^ϵ 表示对 f 加噪声的函数, 并且 $f^\epsilon \in \mathcal{C}([-1, 1])$. L_2 误差会涉及到 L_2 范数的计算, 即(1.2)式, 一致误差会涉及到一致范数的计算, 即

$$\|f\|_\infty = \max_{x \in [-1, 1]} |f(x)|.$$

设 \mathbb{P}_L 表示不超过 L 次的多项式空间, 多项式 $p \in \mathbb{P}_L$ 对函数 f 逼近的最佳逼近误差为

$$E_L(f) := \inf_{p \in \mathbb{P}_L} \|f - p\|_\infty, \quad f \in \mathcal{C}([-1, 1]). \quad (2.1)$$

由 Weierstrass 逼近定理可知,

$$E_L(f) \rightarrow 0 \quad (L \rightarrow \infty). \quad (2.2)$$

设 p^* 表示对函数 f 逼近的 L 次最佳逼近多项式, 那么 $E_L(f) = \|f - p^*\|_\infty$.

2.1.1 L_2 误差

将吉洪诺夫正则化的近似多项式(1.5)考虑为作用在 f 上的算子 $\mathcal{U}_{\lambda, L, N+1} : \mathcal{C}([-1, 1]) \rightarrow L_2([-1, 1])$, 即

$$p_{L, N+1}(x) := \mathcal{U}_{\lambda, L, N+1} f(x) := \frac{1}{1 + \lambda} \sum_{\ell=0}^L \left(\sum_{j=0}^N w_j \tilde{\Phi}_\ell(x_j) f(x_j) \right) \tilde{\Phi}_\ell(x). \quad (2.3)$$

当 $\lambda = 0$ 时, (2.3)式化为 $[-1, 1]$ 上的超插值多项式^[46]

$$\mathcal{U}_{0,L,N+1}f(x) = \sum_{\ell=0}^L \left(\sum_{j=0}^N w_j \tilde{\Phi}_\ell(x_j) f(x_j) \right) \tilde{\Phi}_\ell(x).$$

因此容易得出

$$\mathcal{U}_{\lambda,L,N+1}f = \frac{1}{1+\lambda} \mathcal{U}_{0,L,N+1}f. \quad (2.4)$$

吉洪诺夫正则化的近似多项式(1.5)的 L_2 误差写成如下定理。

定理 2.1.1. ^[1] 设函数 $f \in \mathcal{C}([-1, 1])$, 对 f 加噪声的函数为 $f^\epsilon \in \mathcal{C}([-1, 1])$, $\mathcal{U}_{\lambda,L,N+1}f \in \mathbb{P}_L$ 被(2.3)式定义, 令 $V = \sum_{j=0}^N w_j$, 其中 w_0, \dots, w_N 是高斯求积系数。若 $2L \leq 2N + 1$, 则

$$\|\mathcal{U}_{\lambda,L,N+1}f^\epsilon - f\|_{L_2} \leq \frac{V^{\frac{1}{2}}}{1+\lambda} \|f - f^\epsilon\|_\infty + V^{\frac{1}{2}} \left(1 + \frac{1}{1+\lambda}\right) E_L(f) + \frac{\lambda \|p^*\|_{L_2}}{1+\lambda}, \quad (2.5)$$

其中 p^* 是 f 的 L 次最佳逼近多项式, $E_L(f) = \|f - p^*\|_\infty$ 是最佳逼近的误差。

2.1.2 一致误差

定义

$$\Lambda_{\lambda,L} := \sup_{f \neq 0} \frac{\|\mathcal{U}_{\lambda,L,N+1}f\|_\infty}{\|f\|_\infty} = \sup_{f \neq 0} \frac{\|p_{L,N+1}\|_\infty}{\|f\|_\infty}. \quad (2.6)$$

令 $\lambda = 0$, 则(2.6)式化为 L 次超插值 $\mathcal{U}_{0,L,N+1}$ 的 Lebesgue 常数

$$\Lambda_L := \sup_{f \neq 0} \frac{\|\mathcal{U}_{0,L,N+1}f\|_\infty}{\|f\|_\infty}.$$

由(2.4)式可以得出

$$\Lambda_{\lambda,L} = \frac{1}{1+\lambda} \Lambda_L. \quad (2.7)$$

吉洪诺夫正则化的近似多项式(1.5)的一致误差写成如下定理。

定理 2.1.2. ^[1] 给定函数 $f \in \mathcal{C}([-1, 1])$, 对 f 加噪声的函数为 $f^\epsilon \in \mathcal{C}([-1, 1])$, $\mathcal{U}_{\lambda,L,N+1}f \in \mathbb{P}_L$ 被(2.3)式定义, 令 $V = \sum_{j=0}^N w_j$, 其中 w_0, \dots, w_N 是高斯求积系数。若 $2L \leq 2N + 1$, 则

$$\|\mathcal{U}_{\lambda,L,N+1}f^\epsilon - f\|_\infty \leq \Lambda_{\lambda,L} \|f - f^\epsilon\|_\infty + (1 + \Lambda_{\lambda,L}) E_L(f) + \frac{\lambda \|p^*\|_\infty}{1+\lambda}, \quad (2.8)$$

其中 p^* 是 f 的 L 次最佳逼近多项式, $E_L(f) = \|f - p^*\|_\infty$ 是最佳逼近的误差。

2.2 误差估计的两个推论

在 L_2 误差估计(2.5)和一致误差估计(2.8)中都含有最佳逼近的误差项 $E_L(f)$, 但没有对其进行量化。因此, 本部分会对两种误差估计式中的最佳逼近误差项 $E_L(f)$ 进行量化, 并且也对一致误差估计中的 $\Lambda_{\lambda,L}$ 进行量化, 从而给出更具体的误差估计表达式。

众所周知, 截断的 Chebyshev 展式非常近似最佳逼近多项式^[47]。如果要更严谨地说明这里的“非常近似”, 可以通过 Wang^[48] 的一个计算结果来解释。设 p_L^f 是对 f 的 L 次截断的 Chebyshev 展式, 由文献 [49] 知

$$E_L(f) \leq \|f - p_L^f\|_\infty < \left(4 + \frac{4}{\pi^2} \log L\right) E_L(f). \quad (2.9)$$

由此可见, p_L^f 的一致误差虽然比最佳逼近误差大, 但最多比最佳逼近误差大一个对数因子。Wang^[48] 通过进一步计算表示, 对于 $L \leq 10^6$ 有

$$\|f - p_L^f\|_\infty < 9.6E_L(f).$$

这意味着即使最佳逼近多项式 p^* 和截断的 Chebyshev 展式 p_L^f 的次数达到了一百万次, p^* 的精度也不比 p_L^f 的精度高出一位数。

因此, 在与定理2.1.1和定理2.1.2相同的条件下, 由不等式(2.9), 可以将 L_2 误差(2.5)和一致误差(2.8)分别转化为

$$\|\mathcal{U}_{\lambda,L,N+1}f^\epsilon - f\|_{L_2} \leq \frac{V^{\frac{1}{2}}}{1+\lambda} \|f - f^\epsilon\|_\infty + V^{\frac{1}{2}} \left(1 + \frac{1}{1+\lambda}\right) \|f - p_L^f\|_\infty + \frac{\lambda \|p^*\|_{L_2}}{1+\lambda} \quad (2.10)$$

和

$$\|\mathcal{U}_{\lambda,L,N+1}f^\epsilon - f\|_\infty \leq \Lambda_{\lambda,L} \|f - f^\epsilon\|_\infty + (1 + \Lambda_{\lambda,L}) \|f - p_L^f\|_\infty + \frac{\lambda \|p^*\|_\infty}{1+\lambda}. \quad (2.11)$$

Trefethen 在文献 [50] 中证明了在 $[-1, 1]$ 上 Lipschitz 连续的函数 $f(x)$ 有唯一的 Chebyshev 展式

$$f(x) = \sum_{j=0}^{\infty} \alpha_j T_j(x), \quad \alpha_j = \frac{2}{\pi} \int_{-1}^1 \frac{f(x) T_j(x)}{\sqrt{1-x^2}} dx, \quad j = 0, 1, \dots,$$

其中符号 ' 表示求和的第一项的 1/2, 其余求和项不变, $T_j(x) = \cos(j \cos^{-1} x)$ 表示 j 次的第一类 Chebyshev 多项式。那么, 在多项式空间 \mathbb{P}_L 中, 对函数 $f(x)$ 逼近的多项式就有 $f(x)$ 的 L 次截断 Chebyshev 展式

$$p_L^f(x) = \sum_{j=0}^L {}' \alpha_j T_j(x).$$

下面引理给出 L 次截断 Chebyshev 展式的一致误差。

引理 2.2.1. ^[51] 如果 $f, f', \dots, f^{(k-1)}$ 在 $[-1, 1]$ 上绝对连续, 并且对 $k \geq 1$ 有 $\|f^{(k)}\|_T = M < \infty$ 成立, 那么对每一个 $L \geq k + 1$, 有

$$\|f - p_L^f\|_\infty \leq \frac{2M}{k\pi L(L-1) \cdots (L-k+1)}, \quad (2.12)$$

其中 $\|\cdot\|_T$ 是 Chebyshev 权的 1-范数, 其定义为

$$\|u\|_T := \int_{-1}^1 \frac{|u'(t)|}{\sqrt{1-t^2}} dt.$$

因此, 两不等式(2.10)和(2.11)的右边第二项都可以量化, 因此得出如下的推论。

2.2.1 L_2 误差的推论

推论 2.2.1. 在定理2.1.1和引理2.2.1的条件下,

$$\begin{aligned} \|\mathcal{U}_{\lambda, L, N+1} f^\epsilon - f\|_{L_2} &\leq \frac{V^{\frac{1}{2}}}{1+\lambda} \|f - f^\epsilon\|_\infty + \frac{2MV^{\frac{1}{2}}[1+1/(1+\lambda)]}{k\pi L(L-1) \cdots (L-k+1)} \\ &\quad + \frac{\lambda \|p^*\|_{L_2}}{1+\lambda}. \end{aligned} \quad (2.13)$$

证明 将(2.12)式代入到(2.10)式中, 可以得到(2.13)式。 \square

这样处理 $E_L(f)$ 的方式也参考了文献 [52], 此文献是对重心 Legendre 插值公式的误差进行估计。

当 f 不含噪声, 但是经过正则化时, 由(2.13)式可得此时的 L_2 误差估计为

$$\|\mathcal{U}_{\lambda, L, N+1} f - f\|_{L_2} \leq \frac{2MV^{\frac{1}{2}}[1+1/(1+\lambda)]}{k\pi L(L-1) \cdots (L-k+1)} + \frac{\lambda \|p^*\|_{L_2}}{1+\lambda}.$$

当没有经过正则化, 即正则化参数 $\lambda = 0$, 并且 f 含有噪声时, 由(2.13)式可得此时的 L_2 误差估计为

$$\|\mathcal{U}_{0,L,N+1}f^\epsilon - f\|_{L_2} \leq V^{\frac{1}{2}}\|f - f^\epsilon\|_\infty + \frac{4MV^{\frac{1}{2}}}{k\pi L(L-1)\cdots(L-k+1)}. \quad (2.14)$$

将(2.14)式和前面的(2.13)式进行对比, 容易看出(2.13)式右边第三项消失了, 但导致右边第一项和第二项的值都增加了。因此正则化参数 λ 的选取需要被权衡。

2.2.2 一致误差的推论

一致误差估计(2.11)中含有 $\Lambda_{\lambda,L}$ 的这一项也可以量化。因为 Lebesgue 常数 Λ_L 是在 $[-1, 1]$ 上的超插值的一致算子范数, 因此由文献 [53] 可以知道

$$\Lambda_L \asymp \log L, \quad (2.15)$$

其中 $A(L) \asymp B(L)$ 意味着 $A(L) \lesssim B(L)$ 且 $A(L) \gtrsim B(L)$, 并且 $A(L) \lesssim B(L)$ 表示存在一个与 L 无关的常数 c 使得 $A(L) \leq cB(L)$. 因此结合(2.7)式和(2.15)式有

$$\Lambda_{\lambda,L} \asymp \frac{\log L}{1+\lambda}. \quad (2.16)$$

推论 2.2.2. 在定理2.1.2和引理2.2.1的条件下,

$$\begin{aligned} \|\mathcal{U}_{\lambda,L,N+1}f^\epsilon - f\|_\infty &\leq \frac{c \log L}{1+\lambda} \|f - f^\epsilon\|_\infty + \frac{2M[1 + c \log L / (1 + \lambda)]}{k\pi L(L-1)\cdots(L-k+1)} \\ &\quad + \frac{\lambda \|p^*\|_\infty}{1+\lambda}, \end{aligned} \quad (2.17)$$

其中 c 是一个与 L 无关的常数。

证明 将(2.12)式代入到(2.11)中, 可以得到

$$\begin{aligned} \|\mathcal{U}_{\lambda,L,N+1}f^\epsilon - f\|_\infty &\leq \Lambda_{\lambda,L} \|f - f^\epsilon\|_\infty + \frac{2M(1 + \Lambda_{\lambda,L})}{k\pi L(L-1)\cdots(L-k+1)} \\ &\quad + \frac{\lambda \|p^*\|_\infty}{1+\lambda}. \end{aligned} \quad (2.18)$$

再由(2.16)式和(2.18)式, 可以得到 (2.17)式。 \square

当 f 不含噪声, 但是经过正则化时, 由(2.17)式可得此时的一致误差估计

为

$$\|\mathcal{U}_{\lambda,L,N+1}f - f\|_{\infty} \leq \frac{2M[1 + c \log L/1 + \lambda]}{k\pi L(L-1)\dots(L-k+1)} + \frac{\lambda\|p^*\|_{\infty}}{1 + \lambda},$$

其中 c 是一个与 L 无关的常数。当没有经过正则化，即正则化参数 $\lambda = 0$ ，并且 f 含有噪声时，由(2.17)式可得此时的一致误差估计为

$$\|\mathcal{U}_{0,L,N+1}f^{\epsilon} - f\|_{\infty} \leq c \log L \|f - f^{\epsilon}\|_{\infty} + \frac{2M(1 + c \log L)}{k\pi L(L-1)\dots(L-k+1)}, \quad (2.19)$$

其中 c 是一个与 L 无关的常数。

同样地，将(2.17)式和(2.19)式进行对比，容易看出(2.17)式右边第三项消失了，但导致右边第一项和第二项的值都增加了。因此通过一致误差也再次说明正则化参数 λ 选取的重要性。

2.3 本章小结

本章首先回顾了文献 [1] 中的吉洪诺夫正则化的近似多项式的 L_2 误差估计和一致误差估计。然后对其 L_2 误差界和一致误差界中共有的最佳逼近误差项进行量化。另外，本章还对吉洪诺夫正则化的近似多项式的一致误差界中的 $\Lambda_{\lambda,L}$ 进行量化。然后，本章得出了量化后的吉洪诺夫正则化的近似多项式的 L_2 误差和一致误差的推论。最后通过取 $\lambda = 0$ ，本章得出了此时的 L_2 误差和一致误差的表达式，从而分析说明了吉洪诺夫正则化的近似多项式(1.5)中的正则化参数 λ 选取的重要性。

3. 吉洪诺夫正则化的近似多项式的参数选取

本文采用两种方法选取正则化参数,分别是平衡原则^[3, 28]以及 Brezinski–Rodriguez–Seatzu(BRS) 估计法^[22, 25]。平衡原则的使用需要提前知道噪声水平,但是 BRS 估计法不需要。

3.1 平衡原则

由文献 [11] 第 5.3 节知,吉洪诺夫正则化问题的解通常会带有两部分误差,一部分是由样本数据有噪声引起的,因此需要加入正则项来降低此类误差。正则化参数越大,这部分误差越小。另一部分误差是由逼近引起的,且与样本数据无关,正则化参数越大,这部分误差越大。平衡原则就是在这两种误差之间建立一种平衡的关系。下面针对本文的问题,具体说明并应用平衡原则。

为了使用平衡原则选取(1.5)式中的正则化参数 λ ,需要做如下假设。设 $\epsilon = [\epsilon_0, \epsilon_1, \dots, \epsilon_N]^T \in \mathbb{R}^{N+1}$ 且 $\|\epsilon\|_\infty = \max_i |\epsilon_i|$ 。假设 $\|f - f^\epsilon\|_\infty = \|\epsilon\|_\infty$ 。这一假设意味着采用确定型噪声模型,该模型在 \mathcal{X}_{N+1} 的任意点允许出现最坏的噪声水平。对于最佳逼近误差 $E_L(f)$,假设 L 足够大,由(2.2)式可以确保 $E_L(f) \leq \|\epsilon\|_\infty$ 。如若不然,数据噪声将主要由逼近误差控制,则不需要正则化。在文献 [33] 中对球面多项式的参数选取时,使用平衡原则前也做了类似的两点假设。因此(2.8)式可以做如下改写:

$$\begin{aligned} \|\mathcal{U}_{\lambda,L,N+1}f^\epsilon - f\|_\infty &\leq \Lambda_{\lambda,L}\|\epsilon\|_\infty + (1 + \Lambda_{\lambda,L})\|\epsilon\|_\infty + \frac{\lambda}{1 + \lambda}\|p^*\|_\infty \\ &= d\Lambda_{\lambda,L}\|\epsilon\|_\infty + \frac{\lambda}{1 + \lambda}\|p^*\|_\infty \\ &= \Psi(\lambda) + \Theta(\lambda), \end{aligned} \tag{3.1}$$

其中 d 是与样本数据无关的常数, $\Psi(\lambda) = d\Lambda_{\lambda,L}\|\epsilon\|_\infty$, $\Theta(\lambda) = \frac{\lambda}{1 + \lambda}\|p^*\|_\infty$ 。这

里称 $\Psi(\lambda)$ 为噪声传播误差, 由采样过程被噪声污染引起。由(2.7)式知, $\Psi(\lambda)$ 关于 λ 是单调递减的。称 $\Theta(\lambda)$ 为正则化误差, 由增加的正则项引起。显然 $\Theta(\lambda)$ 关于 λ 是单调递增的。图3.1是 $\Psi(\lambda)$ 和 $\Theta(\lambda)$ 关于 λ 的函数图像, 这两条曲线在 $\lambda = \lambda_*$ 处相交。

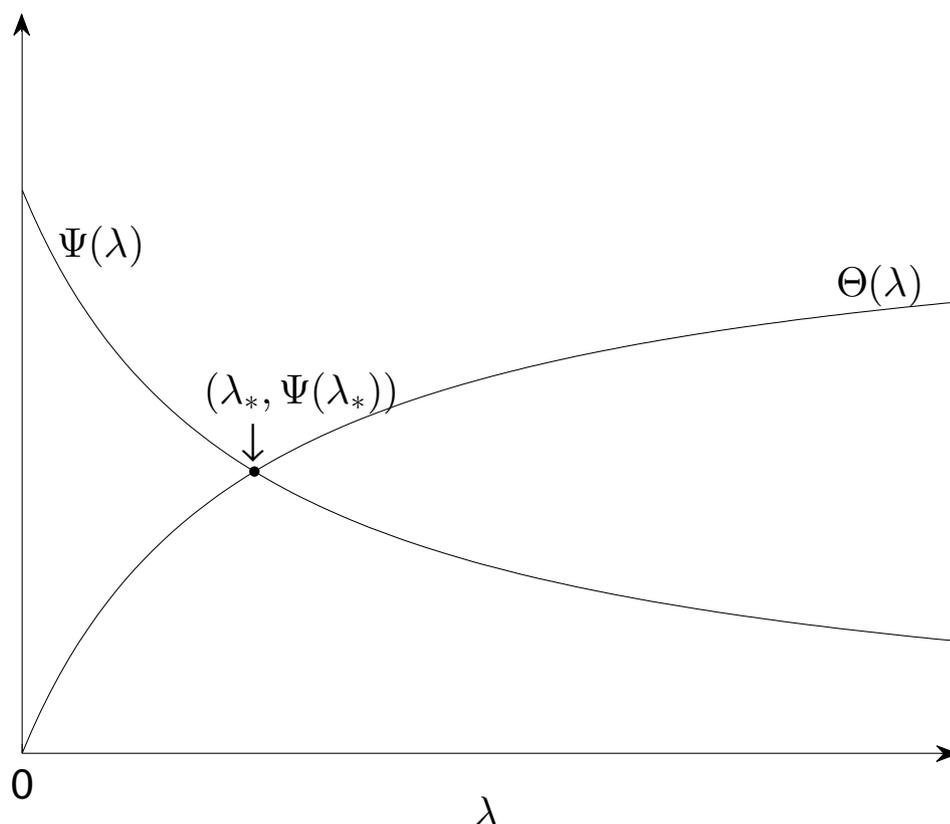


图 3.1: $\Psi(\lambda)$ 和 $\Theta(\lambda)$ 关于 λ 的函数图像

平衡原则就是在这两种误差之间建立一种平衡的关系, 即取 $\lambda = \lambda_*$ 时, 有 $\Psi(\lambda_*) = \Theta(\lambda_*)$. 因此

$$\|\mathcal{U}_{\lambda_*, L, N+1} f^\epsilon - f\|_\infty \leq 2\Psi(\lambda_*) = 2\Theta(\lambda_*).$$

但在许多问题中, Θ 往往是未知的, 不能估计。为了对理想的正则化参数 λ_* 进行估计, 考虑对正则化参数的可能取值进行离散。

取一个有限序列集 $\Delta_K = \{\lambda_i : \lambda_K < \lambda_{K-1} < \cdots < \lambda_0\}$, 其中 K 足够大, 使得理想的正则化参数 λ_* 在两个相邻值之间, 即 $\lambda_{t+1} \leq \lambda_* \leq \lambda_t$. 现任取 Δ_K 中满足 $\lambda_j \leq \lambda_i \leq \lambda_*$ 的两个参数 λ_i 和 λ_j , 则由 $\Psi(\lambda)$ 和 $\Theta(\lambda)$ 的单调

性及(3.1)式得

$$\begin{aligned}
 \|\mathcal{U}_{\lambda_i, L, N+1} f^\epsilon - \mathcal{U}_{\lambda_j, L, N+1} f^\epsilon\|_\infty &\leq \|\mathcal{U}_{\lambda_i, L, N+1} f^\epsilon - f\|_\infty + \|\mathcal{U}_{\lambda_j, L, N+1} f^\epsilon - f\|_\infty \\
 &\leq \Psi(\lambda_i) + \Theta(\lambda_i) + \Psi(\lambda_j) + \Theta(\lambda_j) \\
 &\leq \Psi(\lambda_i) + \Psi(\lambda_j) + 2\Theta(\lambda_i) \\
 &\leq 4\Psi(\lambda_j).
 \end{aligned}$$

因此，平衡原则就是通过如下式子对理想值 λ_* 进行估计：

$$\begin{aligned}
 \lambda_+ = \max\{\lambda_i \in \Delta_K : \|\mathcal{U}_{\lambda_i, L, N+1} f^\epsilon - \mathcal{U}_{\lambda_j, L, N+1} f^\epsilon\|_\infty \leq 4\Psi(\lambda_j), \\
 j = K, K-1, \dots, i\}. \quad (3.2)
 \end{aligned}$$

若违反了(3.2)中的不等式，则所取的参数值接近两曲线的交点 λ_* 。这样选择正则化参数的方式不再依赖 Θ ，并且仍然可以使 λ_+ 充分接近 λ_* 。但是通过(3.2)式来选取参数时，需要在不同 λ_i 处进行大量的比较。由文献 [3] 的第 4.3.1 节可知，这个过程可以通过如下方式进一步简化。

对 Δ_K 中的元素增加一个条件，使得 Δ_K 是一个几何序列的有限集 $\Delta_K = \{\lambda_i = q^i \lambda_0, i = 0, 1, 2, \dots, K\}$ ，其中 $q \in (0, 1)$ 且 K 足够大，那么平衡原则通过如下式子对 λ_* 进行估计：

$$\begin{aligned}
 \lambda_{BP} = \max\{\lambda_z \in \Delta_K : \|\mathcal{U}_{\lambda_j, L, N+1} f^\epsilon - \mathcal{U}_{\lambda_{j+1}, L, N+1} f^\epsilon\|_\infty \leq 4\Psi(\lambda_{j+1}), \\
 j = K-1, K-2, \dots, z+1\}.
 \end{aligned}$$

这样就使得 λ_{BP} 充分接近理想参数值 λ_* 。在应用平衡原则时，从最小的 λ_K 开始，通过 $\lambda_{i-1} = \lambda_i/q$ ， $i = K, K-1, \dots$ ，逐步递增，直到参数 $\lambda_{BP} = \lambda_z$ 第一次满足

$$\|\mathcal{U}_{\lambda_z, L, N+1} f^\epsilon - \mathcal{U}_{\lambda_{z+1}, L, N+1} f^\epsilon\|_\infty > 4\Psi(\lambda_{z+1}) = \mu \Lambda_{\lambda_{z+1}, L} \|\epsilon\|_\infty, \quad (3.3)$$

其中 $\Lambda_{\lambda, L}$ 由(2.6)式所定义， μ 是一个与样本数据无关的设计参数，而平衡原则选择的参数 λ_{BP} 随着样本数据变化而变化。

对于(3.3)式的计算，只需要(2.3)式和(2.16)式。 q 越大， λ 的值增长得越慢。在后面所涉及的平衡原则的数值实验中，都固定 $q = 0.9$ ， $\mu = 0.008$ ， $K = 200$ 和 $\lambda_K = 10^{-3}$ 。

3.2 Brezinski–Rodriguez–Seatzu 估计法

3.2.1 预备知识

Brezinski–Rodriguez–Seatzu 估计法是用发现此方法的三位作者名字命名的一种选取正则化参数的方法，下面都将此方法简称为 BRS 估计法。该方法可用于吉洪诺夫正则化，迭代正则化，截断奇异值分解 (TSVD) 和截断广义奇异值分解 (TGSVD)。其作者对每种正则化给出了具体的参数估计表达式^[25]。本文只涉及到了吉洪诺夫正则化，所以只介绍关于吉洪诺夫正则化的 BRS 估计法。

设吉洪诺夫正则化问题的一般形式为

$$\min_{\mathbf{y} \in \mathbb{R}^n} \{ \|\mathbf{A}\mathbf{y} - \mathbf{b}\|_2^2 + \lambda \|\mathbf{H}\mathbf{y}\|_2^2 \}, \quad (3.4)$$

其中矩阵 $\mathbf{A} \in \mathbb{R}^{m \times n}$ ，向量 $\mathbf{y} \in \mathbb{R}^n$ ，向量 $\mathbf{b} \in \mathbb{R}^m$ 是样本数据 (可能含噪声)， $\lambda > 0$ 为正则化参数， $\mathbf{H} \in \mathbb{R}^{p \times n}$ 是一个正则化矩阵。设 $\|\mathbf{e}\|_2$ 表示最小二乘问题

$$\min_{\mathbf{y} \in \mathbb{R}^n} \|\mathbf{A}\mathbf{y} - \mathbf{b}\|_2$$

的精确解和近似解的误差范数。

BRS 估计法就是用 $\eta_v(\lambda)$ 来估计 $\|\mathbf{e}\|_2$ ，即 $\eta_v(\lambda) \simeq \|\mathbf{e}\|_2$ 。计算解的误差估计 $\eta_v(\lambda)$ ，当 $\eta_v(\lambda)$ 达到最小值时，所取的 λ 值就是 BRS 估计法所选取的最优正则化参数。下面的定理中给出了 $\eta_v(\lambda)$ 的具体表达式，定理的证明见文献 [25]。

定理 3.2.1. ^[25] 对于吉洪诺夫正则化问题(3.4)，设 $\mathcal{N}(\mathbf{A})$ 和 $\mathcal{N}(\mathbf{H})$ 分别表示 \mathbf{A} 和 \mathbf{H} 的零空间，并且满足 $\mathcal{N}(\mathbf{A}) \cap \mathcal{N}(\mathbf{H}) = \{\mathbf{0}\}$ 。固定一个正则化参数 λ ，对应吉洪诺夫正则化问题(3.4)的解设为 \mathbf{y}_λ ，则

$$\eta_v(\lambda) = \lambda^{2-v} \cdot \|\mathbf{b} - \mathbf{A}\mathbf{y}_\lambda\|_2^{v-1} \cdot \|\mathbf{E}\mathbf{y}_\lambda\|_2^{5-v} \cdot \|\mathbf{A}\mathbf{E}\mathbf{y}_\lambda\|_2^{v-3}, \quad v \in \mathbb{R},$$

其中 $\mathbf{E} = \mathbf{H}^T \mathbf{H}$ 。

找到使 $\eta_v(\lambda)$ 最小化的 λ 的方法如下。设 $\xi_1, \xi_2, \dots, \xi_5$ 是 5 个网格点的值，它们将区间 $[\log(\lambda_{\min}), \log(\lambda_{\max})]$ 等距划分成 4 份。对于 $\lambda_i = 10^{\xi_i}, i = 1, \dots, 5$ ，计算出吉洪诺夫正则化问题的解 \mathbf{y}_{λ_i} ，从而得到对应的 $\eta_v(\lambda_i)$ 。然后在这些网格点中找到让 $\eta_v(\lambda)$ 取最小值时的正则化参数 λ_j ，将 λ_j 周围的网格一分为二，

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/288040070024006027>