

小波收缩中统一阈值函数及其偏差、方差与风险分析

赵治栋 潘敏 陈裕泉

(浙江大学生物医学工程与仪器科学学院 杭州 310027)

摘要: 该文建立了小波阈值消噪的统一阈值函数, 推导了统一阈值函数的偏差、方差、风险的明确关系式。利用这些公式研究了参数不同时(以 $u=1,2,\infty$ 为例)统一阈值函数估计的偏差、方差、风险与阈值以及小波系数的关系, 得到了小波统一阈值函数消噪估计的性能, 对小波消噪在工程中应用有重要的理论指导意义。

关键词: 小波收缩, 阈值函数, 偏差估计, 方差估计, 风险估计

中图分类号: TN911.4, 0177.6 **文献标识码:** A **文章编号:** 1009-5896(2005)04-0536-04

Bias, Variance and Risk Analysis of Uniform Threshold Function in Wavelet Shrinkage

Zhao Zhi-dong Pan Min Chen Yu-quan

(Department of Biomedical Engineering and Instruments, Zhejiang University, Hangzhou 310027, China)

Abstract In this paper, the uniform threshold function of waveshrink is build. Computationally efficient formulas for computing bias, variance and risk of uniform threshold function are derived. These formulas provide a new way of understanding how waveshrink works. On the basis of this, the relation of bias, variance and risk of uniform threshold function($u=1,2,\infty$) with threshold value and wavelet coefficients are compared. These comparisons give the performance of waveshrink in finite sample situations.

Key words Wavelet shrinkage, Threshold function, Bias estimation, Variance estimation, Risk estimation

1 引言

小波消噪技术, 目的是从含噪观察信号中估计真实的信号。传统的核平滑和样条平滑技术依赖于待估计的未知信号有关特性的假设, 比如连续性与可微分性, 这些传统技术缺乏空间自适应性, 即估计子随着未知信号的平滑性能做相应调节。由于其多分辨率分析特性, 小波分析非常适用于消噪问题。尤其是当被估计的信号含有一些奇异特性时, 比如突然的跳变, 小波尤其适合^[1-3]。小波消噪的假设是真实信号大部分的小波系数为零, 这是相当合理的一个假设: 小波的消失矩特性及去相关特性保证了真实信号小波系数的稀疏性, 即信号的绝大部分能量集中在少数的小波系数上, 大部分的小波系数为零或接近于零, 而噪声(假设白噪声)的能量均匀分布在每一个小波系数上。Donoho (1994) 开创性地提出了小波阈值收缩消噪技术^[4-6], 开启了小波在统计学上新的研究领域^[7-8]。消噪原理是根据一定的阈值规则向零收缩每一个小波系数达到恢复真实信号的目的, 这个技术的优点在于减少噪声的同时能保持信号的局部奇异性。同时他们提出了软、硬阈值收缩函数, 这两个函数各有优缺点。

本文的主要工作为: (1) 在 Donoho 提出的软硬阈值函

数基础上, 建立统一的阈值函数, 软硬阈值函数只是其两个特例; (2) 推导出统一阈值函数小波收缩估计精确的偏差、方差、风险公式; (3) 分析了参数不同时(以 $u=1,2,\infty$ 为例)统一阈值函数估计的偏差、方差、风险与阈值的关系; (4) 分析了参数不同时(以 $u=1,2,\infty$ 为例)统一阈值函数估计的偏差、方差、风险与小波系数的关系; (5) 通过不同参数的比对得到统一阈值函数估计的性能, 对消噪技术在实际工程中的应用有很大的指导价值。

2 小波消噪模型

假设有 N 个离散点观察信号 y_i 及未知信号 f_i , $i=1,\dots,N$, 则观察模型:

$$y_i = f_i + \varepsilon z_i \quad (1)$$

z_i 为独立同分布(i.i.d)的高斯白噪声, $z_i \sim N(0,1)$, 参数 ε 为噪声强度。目的是从含噪的信号 y 估计出真实的未知的信号 f , 估计性能由估计信号 \hat{f} 与真实信号 f 的均方误差即风险衡量, 定义为

$$R(\hat{f} - f) = E \|\hat{f} - f\|^2 \quad (2)$$

观察模型在小波域可表示为

$$x_i = \theta_i + \varepsilon z_i \quad (3)$$

x_i 与 θ_i 分别为观察信号与真实信号的小波系数, 白噪声 z_i 变换后在小波域仍为白噪声 z_i 。正交小波变换保持 l_2 范数不变, 因此风险式 (2) 可表示为: $R(\hat{\theta} - \theta) = E \|\hat{\theta} - \theta\|^2$ 。

小波消噪主要分为如下步骤: 观察信号正交小波变换、在小波域根据阈值函数对小波系数进行阈值操作、小波系数重构得到消噪后的信号。

3 统一阈值函数及其偏差、方差、风险公式

本文在 Donoho 硬阈值函数与软阈值函数的基础上, 建立了统一的阈值函数:

$$\delta_\lambda^u(x) = x - \lambda^u / (x^{u-1}), \quad u=1, 2, \dots, \infty \quad (4)$$

其中 λ 为阈值。当 u 为偶数时,

$$\delta_\lambda^u(x) = x - xI(|x| \leq \lambda) - \frac{\lambda^u}{x^{u-1}}I(|x| > \lambda) \quad (5)$$

当 u 为奇数时,

$$\delta_\lambda^u(x) = x - xI(|x| \leq \lambda) - \frac{\lambda^u}{x^{u-1}}I(|x| > \lambda)\text{sgn}(x) \quad (6)$$

图 1 为 u 取不同值时统一阈值函数曲线。当 $u=1$ 时, 即为 Donoho 软阈值函数, 当 $u=\infty$ 时, 即为 Donoho 硬阈值函数。由图 1 小波系数较小时, 统一阈值函数接近于软阈值函数; 且 u 越小, 越接近于软阈值函数。在系数较大时, 统一阈值函数接近于硬阈值函数; 且 u 越大, 越接近于硬阈值函数。软阈值估计由于大系数的收缩导致估计有大的偏差。硬阈值估计的不连续性导致有大的方差以及不稳定性: 对信号小的变化比较灵敏。

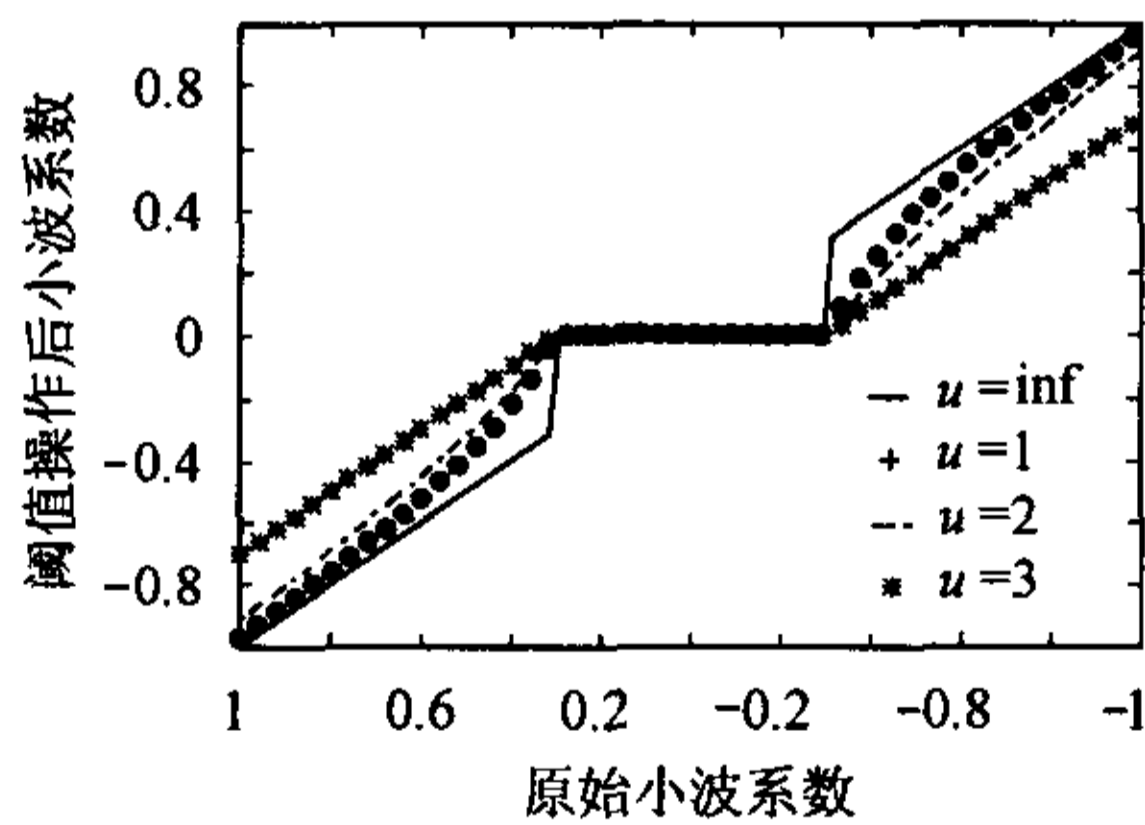


图 1 不同 u 时统一阈值函数曲线

下面给出统一阈值函数估计的期望、偏差、方差与 L_2 风险 (由于篇幅限制具体推导过程从略)。软硬阈值的相应推导公式可见文献[9-12]。下文以下标 H 代表硬阈值, S 代表软阈值, U 代表统一阈值, 即式 (4) 中的参数为 u 。假设 $x \sim N(\theta, 1)$, 定义 θ 的小波统一阈值函数估计的期望 M , 平方偏差 SB , 方差 V 与 L_2 风险 R 。

$$M_U(\lambda, \theta) = E(\delta_\lambda^u(x)) \quad (7)$$

$$SB_U(\lambda, \theta) = (M_U(\lambda, \theta) - \theta)^2 \quad (8)$$

$$V_U(\lambda, \theta) = \text{var}(\delta_\lambda^u(x)) \quad (9)$$

$$R_U(\lambda, \theta) = E(\delta_\lambda^u(x) - \theta)^2 = SB_U(\lambda, \theta) + V_U(\lambda, \theta) \quad (10)$$

ϕ 为标准高斯概率密度函数, Φ 为标准高斯概率分布函数, 由 (8) 式, $SB_U(\lambda, \theta)$ 与 $M_U(\lambda, \theta)$ 存在简单对应关系, 简化起见, 下文仅给出 $M_U(\lambda, \theta)$, $V_U(\lambda, \theta)$, $R_U(\lambda, \theta)$ 公式。

定义

$$A_u(\theta) = \int_\lambda^\infty \frac{\phi(x-\theta) - \phi(x+\theta)}{x^u} dx$$

$$B_u(\theta) = \int_\lambda^\infty \frac{\phi(x-\theta) + \phi(x+\theta)}{x^u} dx$$

则

$$M_H(\lambda, \theta) = \theta + \theta[1 - \Phi(\lambda - \theta) - \Phi(\lambda + \theta)] + \phi(\lambda - \theta) - \phi(\lambda + \theta) \quad (11)$$

$$V_H(\lambda, \theta) = [(\theta^2 + 1)(2 - \Phi(\lambda - \theta) - \Phi(\lambda + \theta)) + (\lambda + \theta)\phi(\lambda - \theta) + (\lambda - \theta)\phi(\lambda + \theta) - M_H(\lambda, \theta)^2] \quad (12)$$

$$R_H(\lambda, \theta) = 1 + (\theta^2 - 1)[\Phi(\lambda - \theta) - \Phi(-\lambda - \theta)] + (\lambda + \theta)\phi(\lambda + \theta) + (\lambda - \theta)\phi(\lambda - \theta) \quad (13)$$

$$M_S(\lambda, \theta) = M_H(\lambda, \theta) - \lambda[\Phi(\lambda + \theta) - \Phi(\lambda - \theta)] \quad (14)$$

$$V_S(\lambda, \theta) = V_H(\lambda, \theta) - \lambda[Q(\lambda, \theta) + Q(\lambda, -\theta)] \quad (15)$$

$$R_S(\lambda, \theta) = 1 + \lambda^2 + (\theta^2 - \lambda^2 - 1)[\Phi(\lambda - \theta) - \Phi(-\lambda - \theta)] - (\lambda - \theta)\phi(\lambda + \theta) - (\lambda + \theta)\phi(\lambda - \theta) \quad (16)$$

$$Q(\lambda, \theta) = [1 + \Phi(\lambda - \theta) - \Phi(\lambda + \theta)] \cdot \{(\theta - \lambda)[1 - \Phi(\lambda - \theta)] + 2\phi(\lambda - \theta)\} \quad (17)$$

$$M_U(\lambda, \theta) = M_H(\lambda, \theta) - \lambda^u A_{u-1}(\theta) \quad (18)$$

$$V_U(\lambda, \theta) = V_H(\lambda, \theta) - 2\lambda^u B_{u-2}(\theta) - \lambda^{2u} A_{u-1}^2(\theta) + \lambda^{2u} B_{2u-2}(\theta) + 2\lambda^u M_H(\lambda, \theta) A_{u-1}(\theta) \quad (19)$$

$$R_U(\lambda, \theta) = V_U(\lambda, \theta) + [M_U(\lambda, \theta) - \theta]^2 = R_H(\lambda, \theta) - 2\lambda^u B_{u-2}(\theta) + \lambda^{2u} B_{2u-2}(\theta) + 2\theta\lambda^u A_{u-1}(\theta) \quad (20)$$

把 $u=1$ 代入式(18)-(20)即可得软阈值的估计式(14)-(16), $u=\infty$ 代入式(18)-(20)即可得硬阈值的估计式(11)-(13)。

4 统一阈值函数估计偏差、方差、风险与阈值的关系

风险 $R(\lambda) = E \|\hat{\theta} - \theta\|^2$, 偏差 $\text{bias}(\lambda) = \|E(\hat{\theta} - \theta)\|^2$ 方差 $\text{var}(\lambda) = E(\|\hat{\theta} - E(\hat{\theta})\|^2)$, 则 $R(\lambda) = \text{bias}(\lambda) + \text{var}(\lambda)$ [13]。若未对小波系数进行阈值操作, 则估计信号为真实信号的无偏估计, $E(\hat{\theta}) = E(Wy) = WE(y) = Wf = \theta$, 但是这个估计方差过大, 风险偏大。若进行阈值操作去掉一部分系数后, 减少了方差, 但是引入了偏差。若进行阈值操作去掉所有的系数, 则方差为零, 但是偏差却增大。因此, 获得最小风险的阈值应是使偏差与方差折衷的阈值。

对于统一阈值函数, 以 $u=2$ 为例, 利用式(18), (19), (20) 研究偏差、方差、风险与阈值 λ 的关系, 其中固定某一 θ 。图 2 从左至右分别为 $\theta=0.2, 1, 4$ 时统一阈值函数偏差、方差、风险与阈值 λ 的关系。

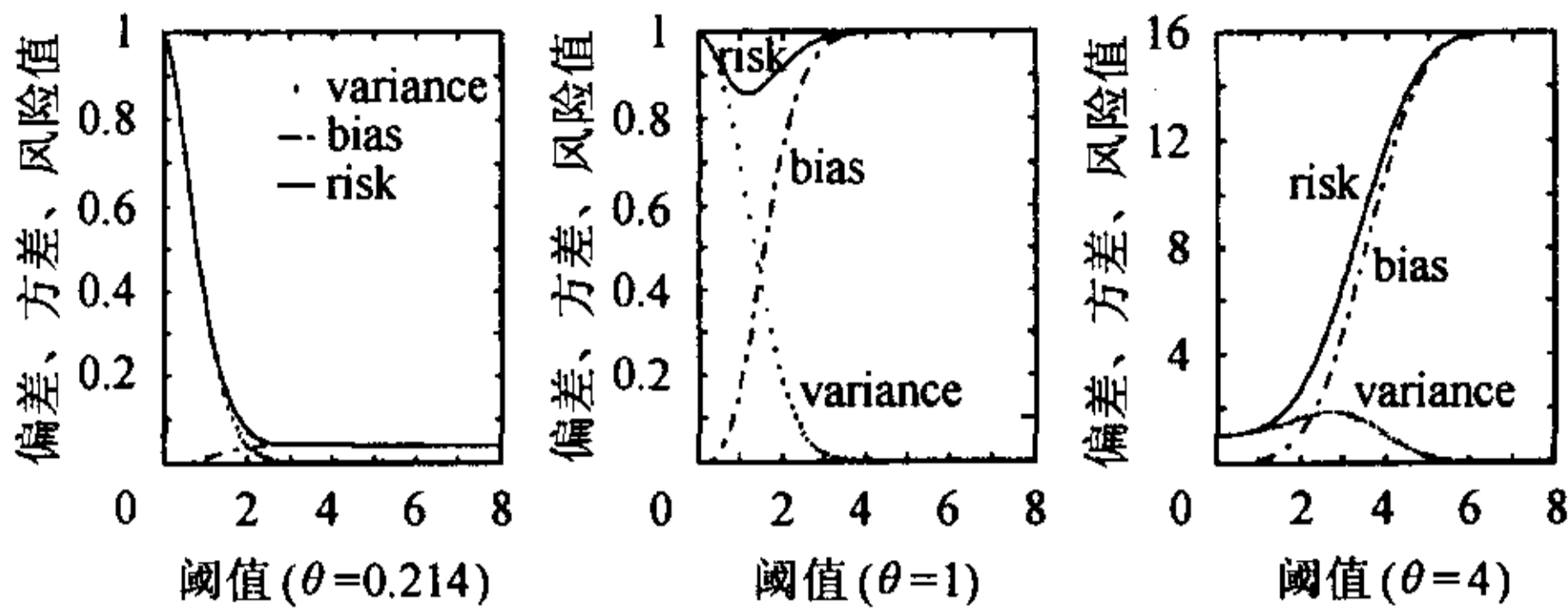


图2 统一阈值函数在不同 θ 时偏差、方差、风险与阈值 λ 的关系

从图中可以看出, (1)统一阈值函数偏差随阈值增加而增加, 偏差来源于阈值操作后抛弃的系数, 阈值越大, 系数被抛弃的可能越大, 偏差越大。方差在系数 θ 较小时, 随阈值增加而减小(图2(a)); 在系数 θ 较大时, 随阈值增加而增加, 若阈值进一步增大, 则方差又会减小, 直至为零(图2(c))。方差来源于阈值操作后保留的系数, 当系数较小时, 阈值越大, 保留的系数越少, 方差越小。如果系数的增加超过了阈值的增加, 则保留的系数随着系数的增大而增大。若系数的增加低于阈值的增加, 则保留的系数随着阈值的增大而减小。最终 $\lim_{\lambda \rightarrow \infty} \text{var}(\lambda) \rightarrow 0$, $\lim_{\lambda \rightarrow \infty} \text{bias}(\lambda) = \|\theta\|^2$, 此时可以认为所有的系数均未保留; 风险则是偏差与方差的和。(2)当系数 θ 较小时风险受偏差影响小, 方差主导着风险, 选择较大阈值较好。 θ 较大时风险受偏差影响大, 方差影响小, 偏差主导着风险, 选择小阈值较好。整体的优化阈值应为偏差与方差的折衷, 从而得到小的风险, 如图2所示。

5 统一阈值函数估计偏差、方差、风险与小波系数关系

根据式(11)~(20), 图(3), (4), (5)分别为硬阈值($u = \infty$)、软阈值($u = 1$)、统一阈值函数(以 $u = 2$ 为例)的偏差、方差、风险与系数 θ 的关系曲线。其中阈值采用 Minimaxi 阈值^[4]。由图可以看出:

(1) 对于固定的阈值 λ , 当 $|\theta|$ 在零附近时, 各阈值函数的偏差和方差较小; 对于硬阈值函数, 当 $|\theta|$ 在 λ 附近时, 方差和偏差均达到最大。 $|\theta|$ 增加时, 偏差和方差逐渐减小, $SB_H \rightarrow 0$, $V_H \rightarrow 1$ 。软阈值的偏差和方差随着 $|\theta|$ 增大逐渐增大, $SB_S \rightarrow \lambda^2$, $V_S \rightarrow 1$ 。统一阈值函数的偏差和方差的变化在 $|\theta|$ 较小时与软阈值函数相似, 在 $|\theta|$ 较大时与硬阈值函数相似。软阈值的偏差一般大于硬阈值的偏差; 软阈值的方差一般比硬阈值小, 统一阈值函数的偏差与方差则是软硬阈值函数的折衷。这也揭示了阈值函数的差异: 硬阈值由于不连续性导致大的方差; 软阈值由于大系数的收缩导致了大的偏差。而统一阈值函数则是软硬阈值的折衷; 其偏差小于软阈值但高于硬阈值; 其方差低于硬阈值但高于软阈值。

(2) 对于一个固定的阈值 λ : 对硬阈值函数, $|\theta|$ 较小在零附近时, 风险较小, $|\theta|$ 在阈值附近时, 风险达到最大, 因为在 $|\theta| \approx \lambda$ 附近, 小波系数有很高的概率为零, 极少数小波系数在阈值附近时, 硬阈值函数比较有效。随着 $|\theta|$ 增加, 风险单调减小。当 $|\theta| \rightarrow \infty$ 时, $R_H(\theta) \rightarrow 1$, 此时小波系数置零的概率很小。对于软阈值函数, $|\theta|$ 较小在零附近时, 风险较小, 随着 $|\theta|$ 增加, 风险单调增加, 当 $|\theta| \rightarrow \infty$ 时, $R_H(\theta) \rightarrow 1 + \lambda^2$, 由于 $|\theta|$ 增大时, 软阈值偏差主导着风险。当大的小波系数集中于近似尺度时, 近似尺度系数不进行阈值操作, 软阈值比较有用。对于统一阈值函数, 在 $|\theta|$ 较大时其风险分布与硬阈值相似, 但是风险变化速率以及峰值均小于硬阈值。 $|\theta|$ 较小时其风险分布与软阈值相似, 从风险上看也是软硬阈值函数的折衷。

(3) 对于不同的阈值 λ 及固定的每一个 θ , 对于所有的阈值函数, 随着 λ 增大, 偏差单调增加。对于软阈值函数, 方差单调减小; 在 $|\theta|$ 较小时, 硬阈值函数和统一阈值函数

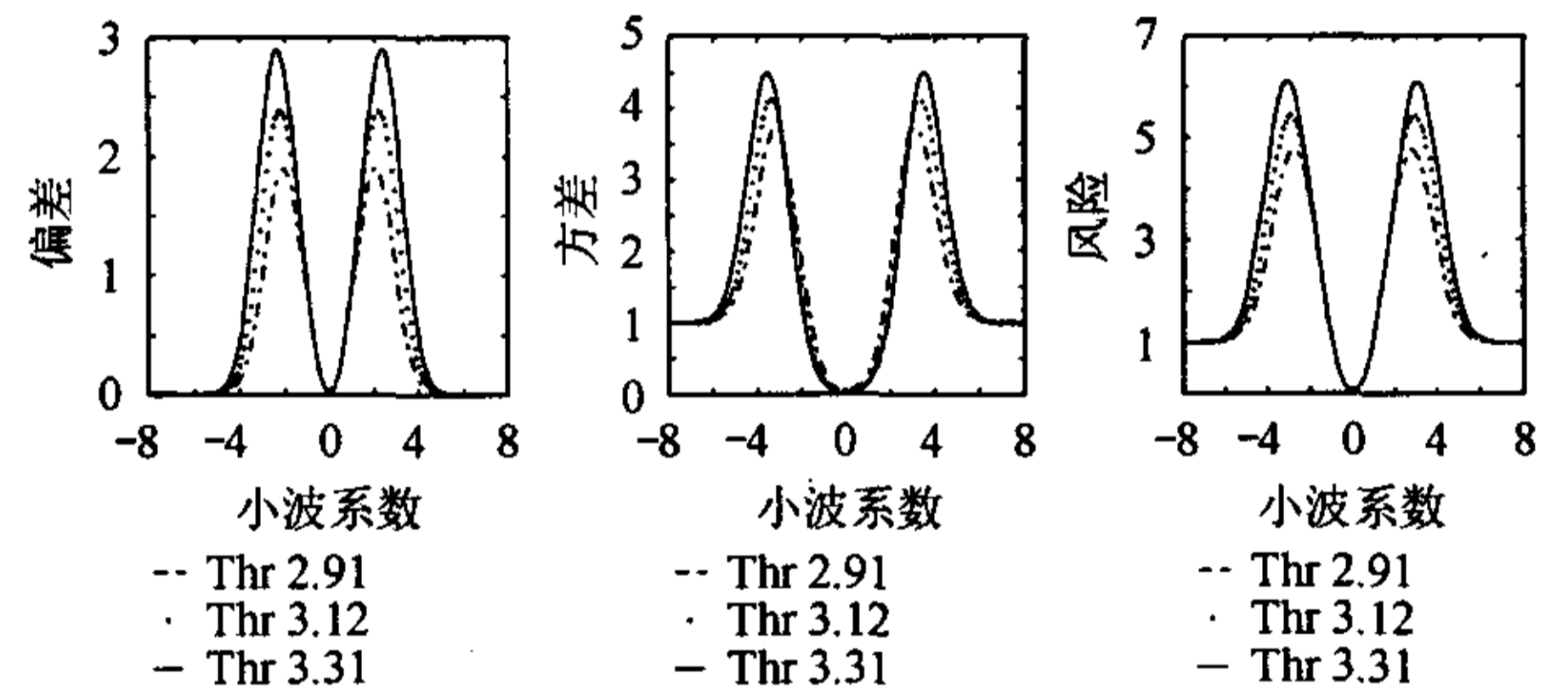


图3 硬阈值函数($u = \infty$)偏差、方差、风险与小波系数关系 (左起偏差、方差、风险)

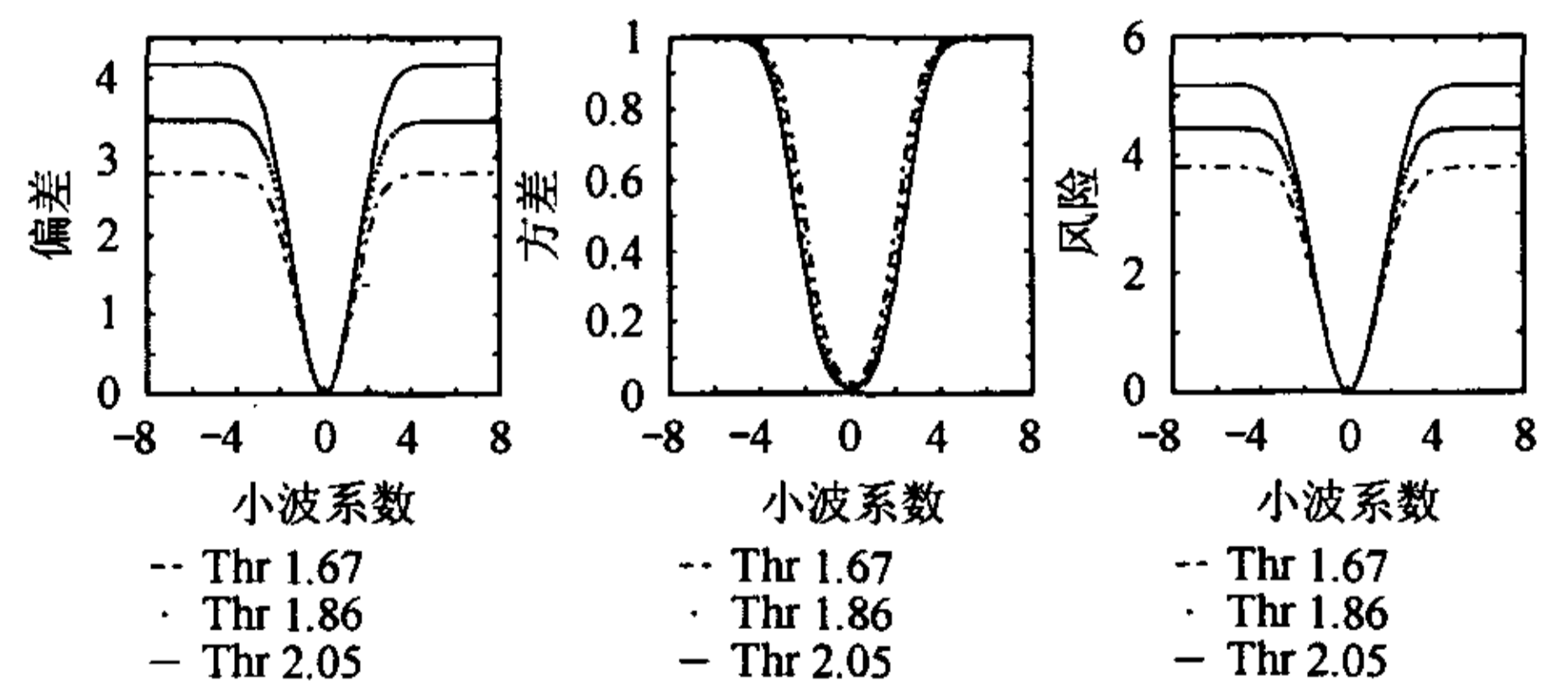


图4 软阈值函数($u = 1$)偏差、方差、风险与小波系数关系 (左起偏差、方差、风险)

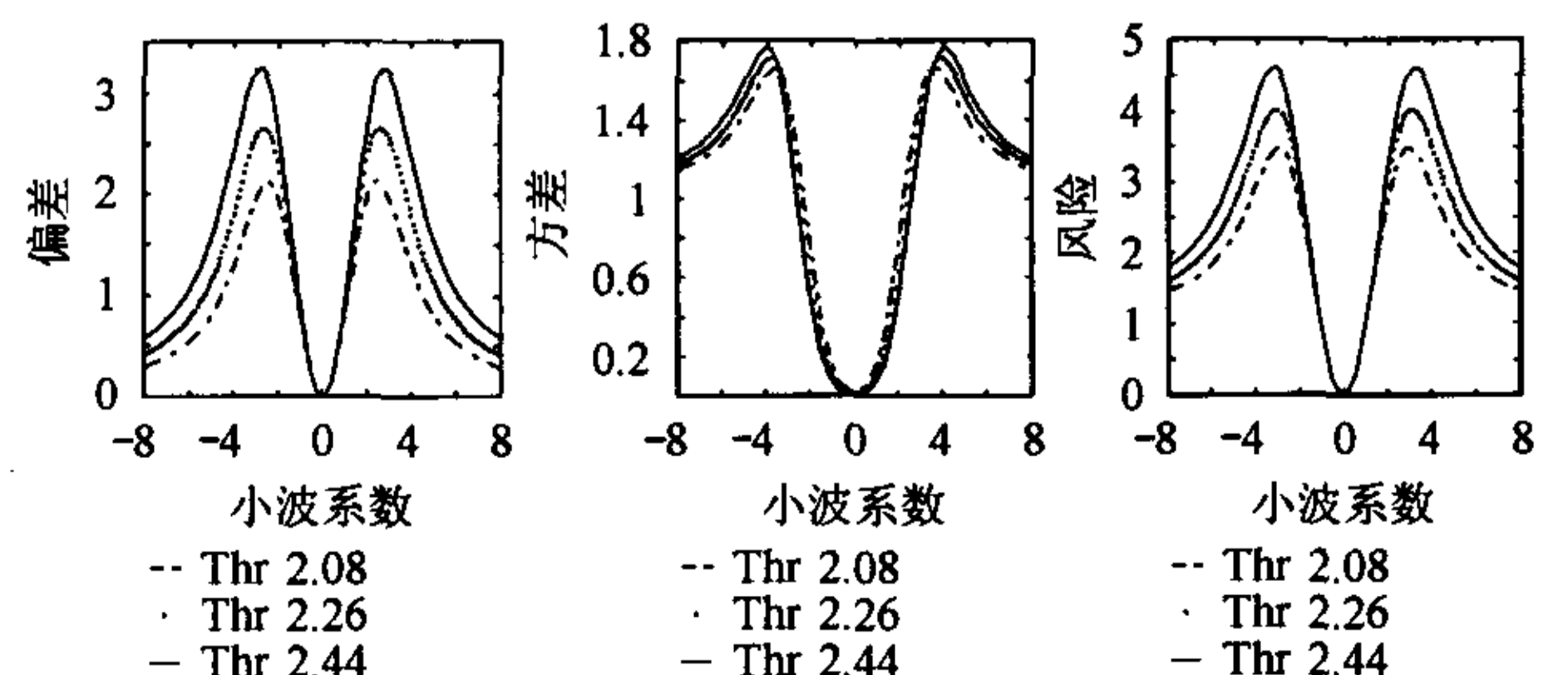


图5 统一阈值函数 ($u = 2$) 偏差、方差、风险与小波系数关系 (左起偏差、方差、风险)

方差单调减小;在 $|\theta|$ 较大时,硬阈值函数和统一阈值函数方差单调增加。对于任何一个阈值函数,风险逐渐增大,除了当系数在原点附近时,风险逐渐减小。

6 结论

小波消噪技术正在各领域信号处理中发挥越来越重要的作用,成为研究的热点,本文从理论上对其性能进行了研究。建立了小波阈值消噪的统一阈值函数,推导了统一阈值函数偏差、方差、风险的公式。研究了小波消噪统一阈值函数(以 $u=1, 2, \infty$ 为例)偏差、方差、风险与阈值和小波系数的关系,得到了统一阈值函数小波估计性能。结果表明,通过恰当的选择统一阈值函数可以获得性能优于既有的软硬阈值函数的阈值函数,一方面避免了硬阈值函数的不稳定性,另一方面减少了软阈值函数较大的估计偏差。本文所得结果对于小波消噪来说有着重要的理论指导意义,已在我们心电、心音等生理信号处理中得到了成功应用。

参考文献

- [1] Mallat S 著,杨力华,戴道清等译.信号处理的小波导引.北京:机械工业出版社,2002:286-327.
- [2] Jansen M. Noise reduction by wavelet thresholding. Springer Verlag, Lecture Notes in Statistics, 2001: 161.
- [3] Taswell C. The what, how, and why of wavelet shrinkage denoising. *Computing in Science & Engineering*, 2000, 2(3): 12-19.
- [4] Donoho D L, Johnstone I. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 1994, 81(3): 425-455.
- [5] Donoho D L. De-noising by soft-thresholding. *IEEE Trans. on Info. Theory*, 1995, 41(3): 612-627.
- [6] Donoho D L, Johnstone I, Kerkacharian G. Wavelet shrinkage: Asymptopia? *J. of the Royal Statist. Soc. Ser. B*, 1995, 57(2): 301-369.
- [7] Antoniadis A. Wavelets in statistics: a review. *J. Ital. Statist. Soc.*, 1997, 6(1): 97-144.
- [8] Abramovich F, Bailey T C, Sapatinas T. Wavelet analysis and its statistical applications. *The Statistician-J. of the Royal Statist. Soc. Ser. D*, 2000, 49(1): 1-29.
- [9] Bruce A G, Gao H Y. WaveShrink:shrinkage functions and thresholds. *SPIE*, 1995, 2569: 270-281.
- [10] Bruce A G, Gao H Y. Understanding waveshrink: variance and bias estimation. *Biometrika*, 1996, 83(4): 727-745.
- [11] Gao H Y. Wavelet shrinkage denoising using the non-negative garrote. *J. Comput. Graph. Statist.*, 1998, 7(4): 469-488.
- [12] Marron J S, Adak S. Exact risk analysis of wavelet regression. *J Comput Graph. Statist.*, 1998, 7(3): 278-309.
- [13] Jansen M. Asymptotic behavior of the minimum mean squared error threshold for noisy wavelet coefficients of piecewise smooth signals. *IEEE Trans. on Signal Proc.*, 2001, 49(6): 1113-1118.

赵治栋: 男, 1976年生, 博士, 主要研究方向为生物医学信号处理.

潘敏: 男, 1969年生, 高级工程师, 主要研究方向为生物传感器、化学传感器、电子技术及系统集成..

陈裕泉: 男, 1944年生, 教授, 博士生导师, 主要研究方向为传感器技术、生物医学信号处理、虚拟仪器学及智能系统等.