Acta Mathematicae Applicatae Sinica, English Series
Vol. 22, No. 3 (2006) 429–436

Acta Mathematicae Applicatae Sinica, English Series

Convergence and Optimality of Adaptive Regularization for Ill-posed Deconvolution Problems in Infinite Spaces

Yan-fei Wang[1], **Qing-hua Ma**[2]

[1]State Key Laboratory of Remote Sensing Science, Jointly Sponsored by the Institute of Remote Sensing Applications, Chinese Academy of Sciences and Beijing Normal University, P.O.Box 9718, Beijing 100101, China (E-mail: yfwang@irsa.ac.cn)

[2]Department of Information Sciences, College of Arts and Science of Beijing Union University, Beijing 100038, China (E-mail: qinghua@ygi.edu.cn)

Abstract The adaptive regularization method is first proposed by Ryzhikov et al. in [6] for the deconvolution in elimination of multiples which appear frequently in geoscience and remote sensing. They have done experiments to show that this method is very effective. This method is better than the Tikhonov regularization in the sense that it is adaptive, i.e., it automatically eliminates the small eigenvalues of the operator when the operator is near singular. In this paper, we give theoretical analysis about the adaptive regularization. We introduce an *a priori* strategy and an *a posteriori* strategy for choosing the regularization parameter, and prove regularities of the adaptive regularization for both strategies. For the former, we show that the order of the convergence rate can approach $O(\|n\|^{\frac{4\nu}{4\nu+1}})$ for some $0 < \nu < 1$, while for the latter, the order of the convergence rate can be at most $O(\|n\|^{\frac{2\nu}{2\nu+1}})$ for some $0 < \nu < 1$.

Keywords Ill-posed problems, adaptive regularization, convergence, regularity
2000 MR Subject Classification 65K10, 65J15

1 Introduction

In remote sensing and geoscience, we are usually required to remove the effects of the unwanted signatures. This is related to a decovolution process of the time series (see [2,4,5,7,9]). For a stationary time series, the traditional method is the Wiener filter. However, in many cases, the records of a time series are highly non-stationary so the Wiener filter is not so effective for reducing instability. Regularization methods have been proved to be a useful tool for suppressing instability (see [1,8,10]). In principle, the deconvolution problem can be considered as a solution of discrete first kind Volterra equation, which is usually solved by regularization. Let us consider a linear system

$$d = Wr + n, \tag{1}$$

where W is a linear operator, r is the input vector and d is the output vector. n is a random vector often containing noise or error. Our purpose is to minimize the noise or error, i.e.,

$$\|n\|^2 \longrightarrow \min.$$

Manuscript received November 29, 2004. Revised February 14, 2006.
Supported by the Chinese Natural Science Foundation of Youth Fund (No. 10501051).

which is equivalent to find a vector $\hat{r}$ such that

$$\mathcal{J} := n^T n = (d - Wr)^T (d - Wr) \longrightarrow \min, \qquad \forall\ r. \tag{2}$$

Let $\nabla \mathcal{J} = 0$. We have

$$W^T W r - W^T d = 0. \tag{3}$$

The above equation can be written as an algebraic equation

$$\Phi r - \widetilde{r} = 0, \tag{4}$$

where $\Phi = W^T W$ is the Fisher matrix or operator and $\widetilde{r} = W^T d$.

(4) is ill-posed in the sense that there exist an infinite number of r' such that $\Phi r' \approx 0$, hence $\Phi(r + r') \approx \Phi r$. Therefore, in computing, the probability of finding $r + r'$ (instead of r) by a computer is 1 almost everywhere. Moreover, Φ is an ill-conditioned operator, so (4) is sensitive to noise. Small perturbations of the data d may lead to significant transfer of the error, so instability occurs.

A proper way to overcome the instability is by the regularization methods. Tikhonov regularization is the well-known one[8]. According to this idea, we add a penalty to $\mathcal{J}$ and have

$$\mathcal{J}_1 = \mathcal{J} + \Delta J = n^T n + \alpha r^T H r, \tag{5}$$

where $H > 0$ is a positive definite matrix, $\alpha > 0$ is the so-called regularization parameter. Differentiating (5) with respect to r, we obtain the discrete Euler equation

$$(\Phi + \alpha H) r = \widetilde{r}. \tag{6}$$

Different H leads to different smoothness of solutions. Clearly, if we choose $H = I$, I is the identity matrix, then (6) is the standard Tikhonov regularization. Ryzhikov and Biryulina[6] chose $H = \Phi^{-1}$ and called the corresponding method adaptive regularization. They had done experiments to show that the adaptive regularization was much better than the standard Tikhonov regularization. We will give a theoretical analysis of their method and prove that the adaptive regularization can approach asymptotical optimality.

We note that the method of Ryzhikov and Bivyulina is formulated in finite spaces. In this paper, we extend their method to infinite spaces and give a theoretical analysis on the convergence and optimality of the adaptive regularization. To be consistent we use the same notation, i.e., (1) is reformulated in an infinite space,

$$W : \mathcal{R} \longrightarrow \mathcal{D}, \qquad Wr = d_n, \tag{7}$$

where $\mathcal{R}$ and $\mathcal{D}$ are both Hilbert spaces which denote the input space (parameter space) and output space (observation space) respectively, $d_n := d_{\text{true}} + n$. In the following, all of the equations (2)–(6) refer to the formulation in infinite spaces. Accordingly, the transpose is replaced by adjoint, i.e., W^T is replaced by W^*, so to others. All of the norms used in the following paragraph are L^2-norm. The following definition is about the spectrum of the adjoint operator Φ, which will be used in later sections.

Definition 1.1[1]. *For any λ, let E_λ be an orthogonal projection operator onto*

$$X_\lambda := span\{u_k : k \in \mathbb{N},\ \beta_k^2 < \lambda\}\ \ (+Null(\Phi),\ if\ \lambda > 0),$$

where $\{\beta_k; u_k\}$ is the eigensystem of Φ and let us define

$$\Phi x = \int \lambda dE_\lambda x, \qquad for\ \ x \in \mathcal{R}.$$

For a given map f, we define

$$\|f(\Phi)x\|^2 = \int f^2(\lambda) d\|E_\lambda x\|^2.$$

The paper is organized as follows. In Section 2, we analyze the convergence properties of the adaptive regularization method when the observation data d_n is error-free, i.e., $d_n \equiv d_{\text{true}}$. In Section 3, the optimality results are obtained under an *a priori* choice of the regularization parameter α. In Section 4, the optimality results are obtained under an *a posteriori* choice of the regularization parameter α. Finally in Section 5, some concluding remarks are given.

2 Convergence of the Adaptive Regularization Method

First, we introduce the concept of the adaptive regularization. The adaptive regularization method is defined as follows. Choosing $H = \Phi^{-1}$ in Euler equation (6) (here we assume that the inverse of Φ exists), we have

$$(\Phi^2 + \alpha I)r = \Phi\widetilde{r}. \tag{8}$$

Define the filter operator $R_\alpha^{\text{adapt}}(\lambda) = \lambda(\lambda^2+\alpha)^{-1}$ and denote the solution of (8) as r_n^α. Then r_n^α can be expressed as

$$r_n^\alpha = R_\alpha^{\text{adapt}}(\Phi)\widetilde{r}. \tag{9}$$

In the ideal case, we can drop the noise n in (1) and have $d_{\text{true}} = Wr^+$, where r^+ denotes the true solution of (1). In this case, the solution of (8) becomes

$$r^\alpha = R_\alpha^{\text{adapt}}(\Phi)\widetilde{r}_{\text{true}}, \tag{10}$$

where $\widetilde{r}_{\text{true}} = W^T d_{\text{true}}$.

It is easy to see that $R_\alpha^{\text{adapt}}(\lambda) \to \lambda^{-1}$ as $\alpha \to 0$. Thus $R_\alpha^{\text{adapt}}(\Phi)$ is an approximation to the inverse of Φ. Defining the Tikhonov filer operator as $R_\alpha^{Tikh}(\lambda) = (\lambda+\alpha)^{-1}$, we have

$$R_\alpha^{\text{adapt}}(\lambda) \longrightarrow 0, \qquad \text{as} \quad \lambda \longrightarrow 0, \tag{11}$$

$$R_\alpha^{Tikh}(\lambda) \longrightarrow \alpha^{-1}, \qquad \text{as} \quad \lambda \longrightarrow 0. \tag{12}$$

This indicates that if the operator Φ is degenerated and has an eigenvalue being null then the adaptive inverse operator eliminates the null-space components. Therefore the adaptive regularization isbetter than the Tikhonov regularization.

Lemma 2.1. $\|r_n^\alpha - r^\alpha\|^2 \le \frac{3\sqrt{3}\|n\|^2}{16\sqrt{\alpha}}$.

Proof.

$$\begin{aligned}
\|r_n^\alpha - r^\alpha\|^2 =& \|R_\alpha^{\text{adapt}}(\Phi)W^T(d-d_{\text{true}})\|^2 \\
=& (WR_\alpha^{\text{adapt}}(\Phi)R_\alpha^{\text{adapt}}(\Phi)W^T(d-d_{\text{true}}), d-d_{\text{true}}) \\
\le& \|R_\alpha^{\text{adapt}}(\Phi)R_\alpha^{\text{adapt}}(\Phi)W^TW\|\|d-d_{\text{true}}\|^2 \\
\le& \|(\Phi^2+\alpha I)^{-2}\Phi^3\|\|n\|^2 \\
\le& \frac{3\sqrt{3}}{16\sqrt{\alpha}}\|n\|^2.
\end{aligned}$$

Lemma 2.2. $r^\alpha := R_\alpha^{\text{adapt}}(\Phi)\widetilde{r}_{\text{true}} \longrightarrow r^+$ *as* $\alpha \longrightarrow 0$, *where,* r^+ *is the true solution.*

Proof. Assume that $\{u_i, v_i; \sigma_i\}$ is the singular system of W, that is,

$$Wu_i = \sigma_i v_i, \qquad W^T v_i = \sigma_i u_i.$$

Then

$$\begin{aligned} R_\alpha^{\text{adapt}}(\Phi)\tilde{r}_{\text{true}} =& R_\alpha^{\text{adapt}}(\Phi)W^T d_{\text{true}} \\ =& \sum_{i=1}^{\infty} \sigma_i R_\alpha^{\text{adapt}}(\sigma_i^2)(d_{\text{true}}, v_i)u_i \\ =& \sum_{i=1}^{\infty} \sigma_i^{-1} \sigma_i^2 R_\alpha^{\text{adapt}}(\sigma_i^2)(d_{\text{true}}, v_i)u_i. \end{aligned}$$

Note that $\lambda R_\alpha^{\text{adapt}}(\lambda) \to 1$ as $\alpha \to 0$, so

$$R_\alpha^{\text{adapt}}(\Phi)\tilde{r}_{\text{true}} \longrightarrow \sum_{i=1}^{\infty} \sigma_i^{-1}(d_{\text{true}}, v_i)u_i = W^+ d_{\text{true}} = r^+.$$

Theorem 2.3. *If $\alpha := \alpha(n) \to 0$ and $\frac{\|n\|^4}{\alpha(n)} \to 0$ as $\|n\| \to 0$, then $r_n^\alpha := R_\alpha^{\text{adapt}}(\Phi)\tilde{r} \to r^+$.*

Proof. We have the following estimate from Lemma 2.1 and Lemma 2.2 that

$$\begin{aligned} \|r_n^\alpha - r^+\| \le& \|r_n^\alpha - r^\alpha\| + \|r^\alpha - r^+\| \\ \le& \frac{\sqrt[4]{27}\|n\|}{4\sqrt[4]{\alpha}} + \|r^\alpha - r^+\|. \end{aligned}$$

Since $\alpha(n) \to 0$ and $\frac{\|n\|^4}{\alpha(n)} \to 0$ as $\|n\| \to 0$, hence $\|r^\alpha - r^+\| \to 0$ (from Lemma 2.2). Thus, $r_n^\alpha \to r^+$.

Remark 2.4. It is well-known that the Tikhonov regularization method is convergent as $\alpha(n) \to 0$ and $\frac{\|n\|^2}{\alpha(n)} \to 0$ (as $\|n\| \to 0$) and that the adaptive regularization method is convergent as long as $\alpha(n) = \|n\|^k$, $k < 4$.

3 Regularity under an *a Priori* Strategy

In this section we analyze the regularity of the adaptive regularization method. First we have the following lemma.

Lemma 3.1. *Assume that $r^+ \neq 0$. Then for $\|n\| > 0$ there exists $\alpha(n)$ such that*

$$\|r^\alpha - r^+\| = \|n\|/\sqrt[4]{\alpha}. \tag{13}$$

In addition, $\alpha(n)$ is strictly monotonically increasing and continuous with

$$\lim_{\|n\| \longrightarrow 0} \alpha(n) = 0, \qquad \lim_{\|n\| \longrightarrow \infty} \alpha(n) = \infty.$$

Proof. Noting that $Wr^+ = d_{\text{true}}$, we have

$$r^\alpha - r^+ = -\alpha(\Phi^2 + \alpha I)^{-1} r^+.$$

By Definition 1.1, we denote E_λ as the spectral family of Φ. Then

$$\sqrt{\alpha}\|r^\alpha - r^+\|^2 = \int_0^\infty \frac{\sqrt{\alpha}\alpha^2}{(\lambda^2+\alpha)^2} d\|E_\lambda r^+\|^2.$$

Defining

$$\phi(\alpha) := \int_0^\infty \frac{\sqrt{\alpha}\alpha^2}{(\lambda^2+\alpha)^2} d\|E_\lambda r^+\|^2 - \|n\|^2,$$

and noting that $r^+ \in Null(W)^\perp$ and $r^+ \neq 0$, we then have $\phi(\alpha)$ is monotonically increasing and continuous with

$$\lim_{\alpha\longrightarrow 0}\phi(\alpha) = -\|n\|^2, \qquad \lim_{\alpha\longrightarrow\infty}\phi(\alpha) = \infty.$$

This completes the proof.

If we further have some *a priori* knowledge of the solution, say, $r^+ \in \text{Range}(\Phi^\nu)$, $0<\nu<1$, we have the following estimate of the order of convergence.

Theorem 3.2. *Assume that* $r^+ \in Range(\Phi^\nu)$, $0<\nu<1$. *If* $\int_0^\mu d\|E_\lambda r^+\|^2 = O(\mu^{2\nu})$, *then* $\|r^+ - r^\alpha\| = O(\alpha^{\frac{\nu}{2}})$.

Proof. First we have the estimate

$$\begin{aligned}\|r^\alpha - r^+\|^2 &= \int_0^\infty \frac{\alpha^2}{(\alpha+\lambda^2)^2} d\|E_\lambda r^+\|^2 \\ &= \int_0^{\sqrt{\alpha}} \frac{\alpha^2}{(\alpha+\lambda^2)^2} d\|E_\lambda r^+\|^2 + \int_{\sqrt{\alpha}}^\infty \frac{\alpha^2}{(\alpha+\lambda^2)^2} d\|E_\lambda r^+\|^2.\end{aligned}$$

Since

$$\frac{1}{4}\int_0^{\sqrt{\alpha}} d\|E_\lambda r^+\|^2 \le \int_0^{\sqrt{\alpha}} \frac{\alpha^2}{(\alpha+\lambda^2)^2} d\|E_\lambda r^+\|^2 \le \int_0^{\sqrt{\alpha}} d\|E_\lambda x^+\|^2,$$

$$0 \le \int_{\sqrt{\alpha}}^\infty \frac{\alpha^2}{(\alpha+\lambda^2)^2} d\|E_\lambda r^+\|^2 \le \frac{1}{4}\int_{\sqrt{\alpha}}^\infty d\|E_\lambda r^+\|^2,$$

and $\int_0^\mu d\|E_\lambda r^+\|^2 = O(\mu^{2\nu})$, then

$$\|r^\alpha - r^+\| = O(\alpha^{\frac{\nu}{2}}).$$

Note that in applications the data is usually contaminated by noise. For example, in remote sensing, the data obtained is usually perturbed due to long-term exposure through the atmosphere where turbulence in the atmosphere gives rise to random variations in the refractive index. Therefore, the data d contains noise. Suppose the noise in d can be controlled by the noise level $\|n\|$, that is,

$$\|d - d_{\text{true}}\| \le \|n\|.$$

Then we can show that the adaptive regularization can really approach the regularity.

Theorem 3.3. *Under the condition of theorem 3.2, we have*

$$\sup\{\inf_{\alpha>0}\|r^+ - r_n^\alpha\| : \|d - d_{\text{true}}\| \le \|n\|\} = O(\|n\|^{\frac{4\nu}{4\nu+1}}). \tag{14}$$

Proof. Suppose $r^+ \neq 0$ (the result is obvious as $r^+ = 0$). From Lemma 2.1 and the triangle inequality, we can show that for all $\alpha > 0$ and d, $\|d - d_{\text{true}}\| \le \|n\|$,

$$\|r_n^\alpha - r^+\| \le \|r^\alpha - r^+\| + \frac{\sqrt[4]{27}\|n\|}{4\sqrt[4]{\alpha}}. \tag{15}$$

Now let $\alpha = \alpha(n)$ be as in Lemma 3.1 and define

$$c_n := \frac{\|r^\alpha - r^+\|}{\alpha^\nu}.$$

Then

$$\alpha^{4\nu} c_n^4 = \frac{\|n\|^4}{\alpha}$$

and α can be expressed as

$$\alpha = (\|n\| c_n^{-1})^{\frac{4}{4\nu+1}}.$$

Hence from (15) we have

$$\sup\{\inf_{\alpha>0} \|r^+ - r_n^\alpha\| : \|d - d_{\text{true}}\| \le \|n\|\} = O\Big(\|r^\alpha - r^+\| + \frac{\sqrt[4]{27}\|n\|}{4\sqrt[4]{\alpha}}\Big)$$
$$= O\Big(\frac{\|n\|}{\sqrt[4]{\alpha}}\Big) = O\big(\|n\|^{\frac{4\nu}{4\nu+1}} c_n^{\frac{1}{4\nu+1}}\big).$$

4 Regularity under an *a Posteriori* Strategy

From the former section we know that the optimal order of convergence is obtained if the choice of $\alpha = \alpha(n)$ is in an *a priori* way, i.e., $\alpha = (\|n\| c_n^{-1})^{\frac{4}{4\nu+1}}$. However this is not applicable in practice. Practically, an *a posteriori* way will be workable. We use the widely used Morozov's discrepancy principle, that is, $\alpha = \alpha(n)$ *should be chosen as*

$$\alpha(n) := \sup\{\alpha > 0 : \|d - W r_n^\alpha\| \le \tau \|n\|\} \tag{16}$$

with $\tau > 1$ *being another parameter.*

Denoting $q_\alpha(\lambda) = 1 - \lambda R_\alpha^{\text{adapt}}(\lambda)$, we find that $q_\alpha(\lambda) \le \beta < 1$, where β is the supremum of $q_\alpha(\lambda)$. We also note that $q_\alpha(\lambda) \to 0$ as $\alpha \to 0$, so

$$\begin{aligned}\|d - W r_n^\alpha\| &= \|d - W(\Phi^2 + \alpha I)^{-1} \Phi \tilde{r}\| \\ &= \|(I - R_\alpha^{\text{adapt}}(WW^T)(WW^T))d\| \le \epsilon \quad \text{as} \quad \alpha \longrightarrow 0,\end{aligned}$$

where ϵ is a small positive number. This shows that (16) can be satisfied.

In the following we show that the adaptive regularization method with the discrepancy principle as the stopping rule can approach the regularity. The results rely on the property of the function

$$f_\alpha(\lambda) := \lambda^{\nu+\frac{1}{2}} q_{2\alpha}(\lambda), \qquad 0 < \nu < 1.$$

A easy calculation shows that $f_\alpha(\lambda)$ can be maximized if and only if $\lambda = \lambda^*$, where

$$\lambda^* = (C_\nu \alpha)^{\frac{1}{2}}$$

with $C_\nu = \frac{4\nu+2}{3-2\nu}$. So the maximum value of $f_\alpha(\lambda)$ is

$$f_\alpha^{\max}(\lambda) = D_\nu \alpha^{\frac{2\nu+1}{4}} \tag{17}$$

with $D_\nu = \frac{2C_\nu^{\frac{2\nu+1}{4}}}{C_\nu+2}$.

Now we give the regularity result:

Theorem 4.1. *Assume that $r^+ \in Range(\Phi^\nu)$, $0 < \nu < 1$. If $\alpha(n)$ is chosen by the above discrepancy principle, then*

$$\|r_n^\alpha - r^+\| = O\big(\|n\|^{\frac{2\nu}{2\nu+1}}\big). \tag{18}$$

Proof. Let $r^+ = \Phi^\nu\omega$, $\omega \in \mathrm{Range}(\Phi^\nu)$, $0 < \nu < 1$. Then by the triangular inequality

$$\|r^+ - r_n^\alpha\| \le \|r_n^\alpha - r^\alpha\| + \|r^\alpha - r^+\|, \tag{19}$$

we need to estimate the upper bounds of $\|r_n^\alpha - r^\alpha\|$ and $\|r^\alpha - r^+\|$.

By Hölder's inequality, we have

$$\begin{aligned}
\|r^\alpha - r^+\| =& \|(\Phi^2 + \alpha I)\Phi\tilde{r}_{\text{true}} - r^+\| \\
=& \|(\Phi^2 + \alpha I)\Phi^2 r^+ - r^+\| \\
=& \|\Phi^\nu q_\alpha(\Phi)\omega\| \\
\le& \|q_\alpha(\Phi)\omega\|^{\frac{1}{2\nu+1}} \|\Phi^{\frac{1}{2}} q_\alpha(\Phi)\Phi^\nu\omega\|^{\frac{2\nu}{2\nu+1}} \\
=& \|q_\alpha(\Phi)\omega\|^{\frac{1}{2\nu+1}} \|Wr^\alpha - d_{\text{true}}\|^{\frac{2\nu}{2\nu+1}} \\
\le& (\beta\|\omega\|)^{\frac{1}{2\nu+1}} \|Wr^\alpha - d_{\text{true}}\|^{\frac{2\nu}{2\nu+1}}.
\end{aligned} \tag{20}$$

Recalling that

$$\begin{aligned}
\|Wr^\alpha - d_{\text{true}}\| =& \|W(r^\alpha - r_n^\alpha) - (d_{\text{true}} - d) + (Wr_n^\alpha - d)\| \\
=& \|WR_\alpha^{\text{adapt}}(\Phi)W^T(d - dtrue) - (d - d_{\text{true}}) + (Wr_n^\alpha - d)\| \\
\le& \|q_\alpha(WW^T)(d - d_{\text{true}})\| + \|Wr_n^\alpha - d\| \\
\le& \beta\|n\| + \tau\|n\| \\
=& (\beta + \tau)\|n\|,
\end{aligned} \tag{21}$$

we obtain

$$\begin{aligned}
\|r^\alpha - r^+\| \le& \big(\beta\|\omega\|\big)^{\frac{1}{2\nu+1}} \big((\beta+\tau)\|n\|\big)^{\frac{2\nu}{2\nu+1}} \\
\le& O\big(\|\omega\|^{\frac{1}{2\nu+1}} \|n\|^{\frac{2\nu}{2\nu+1}}\big).
\end{aligned} \tag{22}$$

Now let $\gamma = 2\alpha$. By the discrepancy principle (16), we have

$$\|q_\gamma(WW^T)d\| = \|d - Wr_n^\gamma\| > \tau\|n\|.$$

By the above inequality and (20), we have

$$\|d_{\text{true}} - Wr^\gamma\| \ge \tau\|n\| - \beta\|n\| = (\tau - \beta)\|n\|,$$

where $\tau - \beta > 0$ since $\tau > 1$, $0 < \beta < 1$. Thus

$$\|n\| \le \frac{1}{\tau - \beta}\|d_{\text{true}} - Wr^\gamma\| = \frac{1}{\tau - \beta}\|\Phi^{\frac{1}{2}} q_\gamma(WW^T)\Phi^\nu\omega\|.$$

By (17), the above inequality gives

$$\|n\| \le \frac{1}{\tau - \beta} D_\nu \alpha^{\frac{2\nu+1}{4}} \|\omega\|. \tag{23}$$

This indicates that

$$\alpha \ge E_{\nu,\tau,\beta}\|\omega\|^{-\frac{4}{2\nu+1}} \|n\|^{\frac{4}{2\nu+1}}, \tag{24}$$

where $E_{\nu,\tau,\beta}$ is a constant with respect to ν, τ and β.

Recalling Lemma 2.1, we have

$$\|r_n^\alpha - r^\alpha\| \le \frac{\sqrt[4]{27}}{4}\|n\|\alpha^{-\frac{1}{4}}. \tag{25}$$

By (24) and (25), we obtain that

$$\|r_n^\alpha - r^\alpha\| \le \frac{\sqrt[4]{27}}{4}(E_{\nu,\tau,\beta})^{-\frac{1}{4}}\|\omega\|^{\frac{1}{2\nu+1}}\|n\|^{\frac{2\nu}{2\nu+1}}. \tag{26}$$

By (19), (22) and (26) we obtain (18).

5 Conclusion

This paper establishes the convergence and regularity results of the adaptive regularization method. We find that for the *a priori* choice of the regularization parameter, the rate of convergence is $O(\|n\|^{\frac{4\nu}{4\nu+1}})$ for some $0 < \nu < 1$. However, for the *a posteriori* choice of the regularization parameter α, the optimal convergence rate is $O(\|n\|^{\frac{2\nu}{2\nu+1}})$. This indicates that, sometimes, if some *a priori* knowledge-based information is known in advance, then we can get better results. This phenomenon is particular useful in remote sensing (see [3]), where the *a priori* information is based on the historical data constraints on the solution, i.e., the retrieved land surface parameters. Ryzhikov and Biryulina[6] pointed out that this method was very important in deconvolution problems which arised frequently in geophysical sciences.

References

[1] Douglas, M.B., Donald, G.W. Nonlinear regression analysis and its applications. John Wiley & Sons, Inc., 1988

[2] Engl, H.W., Hanke, M., Neubauer. Regularization of inverse problems. Kluwer, Dordrecht, 1996

[3] Li, X.W., Gao, F., Wang, J.D., Strahler, A. A priori knowledge accumulation and its application to linear BRDF model inversion. *Journal of Geophysical Research*, 106: 11925–11935 (2001)

[4] Richards, J.A., Jia, X.P. Remote sensing digital image analysis. Springer-Verlag, Berlin, 1999

[5] Robinson, E., Treitel, S. Geophysical signal analysis. Prentice-Hall, Englewood Cliffs, 1980

[6] Ryzhikov, G.A., Biryulina, M.S. Sharp deconvolution in elimination of multiples. 68th Ann. Internat. Mtg. Soc. Expl. Geophys., 1983–1986 (1998)

[7] Seber, G.A.F., Wild, C.J. Nonliear regression. Wiley Interscience, New York, 1989

[8] Tikhonov, A.N., Arsenin, V.Y. Solutions of Ill-posed problems. Wiley, New York, 1977

[9] Webster, G., Levin, F. Deconvolution. Geophysics Reprint Series No.1, Society of Exploration Geophysicists, Tulsa, 1978

[10] Xiao, T.Y., Yu, S.G., Wang, Y.F. Numerical methods for inverse problems. Science Press, Beijing, 2003