UNIVERSITÄT KAISERSLAUTERN

REGULARIZED APPROXIMATION METHODS WITH PERTURBATIONS FOR ILL-POSED OPERATOR EQUATIONS

M. Thamban Nair and Eberhard Schock

Preprint Nr. 231

FACHBEREICH MATHEMATIK

REGULARIZED APPROXIMATION METHODS WITH PERTURBATIONS FOR ILL–POSED OPERATOR EQUATIONS

M. Thamban Nair and Eberhard Schock

Preprint Nr. 231

UNIVERSITÄT KAISERSLAUTERN
Fachbereich Mathematik
Erwin–Schrödinger–Straße
6750 Kaiserslautern

November 1992

REGULARIZED APPROXIMATION METHODS WITH PERTURBATIONS FOR ILL-POSED OPERATOR EQUATIONS

M. Thamban Nair
Department of Mathematics
Goa University, Goa, India

Eberhard Schock
Fachbereich Mathematik
Universität Kaiserslautern, Kaiserslautern, Germany

ABSTRACT

We are concerned with a parameter choice strategy for the Tikhonov regularization

$$(\tilde{A}+\alpha I)\tilde{x} = T^*\tilde{y} + w$$

where $\tilde{A}$ is a (not necessarily selfadjoint) approximation of T^*T and $T^*\tilde{y}+w$ is a perturbed form of the (not exactly computed) term T^*y. We give conditions for convergence and optimal convergence rates.

INTRODUCTION

We are concerned with the problem of finding approximations to the minimal norm least–square solution $\hat{x}$ of the operator equation

$$Tx = y\ , \qquad (1)$$

where T is a bounded linear operator between Hilbert spaces X and Y and y belongs to $D(T^+) := R(T) + R(T)^\perp$, the domain of the Moore–Penrose inverse T^+ of T. It is well known [6] that if the range R(T) of T is not closed, then the operator T^+ (which associates $y \in D(T^+)$ the unique element $\hat{x} := T^+y$ of minimal norm such that $\|T\hat{x}-y\| = \inf\{\|Tx-y\|: x\in X\}$)

is not continuous, and hence the problem of solving (1) for $\hat{x}$ is ill–posed. So, regularization methods are employed to find approximations to $\hat{x}$. A well known such method is the Tikhonov regularization in which a well–posed equation

$$(T^*T+\alpha I)x_\alpha = T^*y \tag{2}$$

is solved for each $\alpha > 0$. It can be easily seen that if $y^\delta \in Y$ with $\|y-y^\delta\| \leq \delta$, is available instead of y, then the solution of

$$(\alpha I+T^*T)x_\alpha^\delta = T^*y^\delta \tag{3}$$

satisfies

$$\|x_\alpha - x_\alpha^\delta\| \leq \frac{\delta}{\sqrt{\alpha}} \longrightarrow 0 \text{ as } \delta \longrightarrow 0.$$

In order that x_α^δ to be an approximation to $\hat{x}$, we must have $x_\alpha^\delta \longrightarrow \hat{x}$ as $\delta \longrightarrow 0$ and $\alpha \longrightarrow 0$. It is known (c.f. Groetsch [6]) that

$$\|\hat{x}-x_\alpha\| \longrightarrow 0 \text{ as } \alpha \longrightarrow 0\ . \tag{4}$$

Therefore a sufficient condition for $x_\alpha^\delta \longrightarrow \hat{x}$ is to choose $\alpha = \alpha(\delta)$ such that $\delta/\sqrt{\alpha(\delta)} \longrightarrow 0$ as $\delta \longrightarrow 0$. Next question is how fast the convergence is. This question is generally tackled under certain "smoothness" assumptions on the unknown $\hat{x}$. If $\hat{x} \in R((T^*T)^\nu)$, $0 < \nu \leq 1$, then it is known [6] that

$$\|\hat{x}-x_\alpha\| = 0(\alpha^\nu)\ , \tag{5}$$

so that an a priori parameter choice $\alpha = \alpha(\delta)$ for max $\{\alpha^\nu, \delta/\sqrt{\alpha}\}$ to be minimum is $\alpha(\delta) = c\delta^{2/(2\nu+1)}$ for some constant $c > 0$. With this choice of α, we have

$$\|\hat{x}-x_\alpha^\delta\| \leq \|\hat{x}-x_\alpha\| + \frac{\delta}{\sqrt{\alpha}} = 0(\delta^{2\nu/(2\nu+1)}). \tag{6}$$

It is also known (Schock [11]) that the rate $0(\delta^{2\nu/(2\nu+1)})$ is optimal, in the sense that it can not be replaced (in general) by $o(\delta^{2\nu/(2\nu+1)})$ for $\hat{x} \in R((T^*T)^\nu)$, unless $\hat{x} = 0$. In a posteriori parameter choice strategies, the parameter $\alpha = \alpha(\delta)$ is choosen during the course of computation of x^{δ}_{α}. Many works in the literature devoted to this (c.f. Groetsch [5], Schock [12], Engl [1], Engl and Neubauer [3], Gfrerer [4], Nair [9], and the references there in.)

In all the above refered works it is presumed that T^*y^{δ} is computed exactly. But, due to the ill–posed nature of the operator T it is very likely that one actually deals with a perturbed quantity $T^*y^{\delta}+w$ with $\|w\|$ small instead of T^*y^{δ} , so that $T^*y^{\delta}+w$ does not belong to the range of T^*. Such situation in the finite dimensional frame work has been considered by Schock [13]. This consideration would be of very importance when one solves equation (3) numerically. For example the operator $A := T^*T$ and T^* may be approximated by sequences of operators (A_n) and (T^*_n) respectively, and in that case one generally considers the problem to be solved is

$$(A_n+\alpha I)x = T^*_n y^{\delta} ;$$

but the actual problem at hand is of the form

$$(A_n+\alpha I)x = T^*_n y^{\delta} +w_n$$

where w_n is some additional error due to computation involved in T^*y^{δ}. The main purpose of this paper is to address this issue. We propose a priori and a posteriori parameter choice strategies for choosing α taking into consideration such inaccuracies due to computation. In particular we show that if A_n and z^{δ}_n are "good" approximations to A and T^*y^{δ} respectively, then the optimal theoretical rate in (6) can be achieved.

THE METHOD AND ITS CONVERGENCE

Let $y^{\delta} \in Y$ with $\|y-y^{\delta}\| \le \delta$, (z^{δ}_n) in X such that

$$\|z^{\delta}_n-T^*y^{\delta}\| \le \eta_{n,\delta} \longrightarrow 0 \quad \text{as} \quad n \to \infty$$

for each $\delta > 0$, and (A_n) be a sequence of bounded linear operators on X such that

$$\|A-A_n\| \leq \varepsilon_n \longrightarrow 0 \text{ as } n \longrightarrow \infty ,$$

where $A := T^*T$. Since A is a non–negative and self–adjoint operator, it follows that for each $n=1,2,\ldots$, $\alpha \geq \varepsilon_n/c_1$ for some constant c_1 such that $0 < c_1 < 1$ implies the existence of $(A_n+\alpha I)^{-1}$ as a bounded linear operator and $\|(A_n+\alpha I)^{-1}\| \leq 1/(1-c_1)\alpha$. It is to be remarked that if A_n is non–negative and self–adjoint, then the above conclusions hold for all $\alpha>0$. In this context one recalls that in projection methods ([8],[3]) for compact T, one has non–negative self–adjoint operators $A_n = P_nT^*TP_n$ where (P_n) is a sequence of orthogonal projections such that $P_nx \longrightarrow x$ as $n \longrightarrow \infty$ for all $x \in X$; and in a degenerate kernel method for integral equations of the first kind considered by Groetsch [7], non–negative self–adjoint operators A_n are obtained by approximating the kernel of the integral operator $A = T^*T$ by convergent quadrature rules. In both these cases one also has $\|A-A_n\| \longrightarrow 0$. But the consideration of a general A_n, as has been done recently by Nair [10], is again important in computational point of view, because due to truncation errors etc., one actually may not be dealing with a non–negative self–adjoint operator.

We consider the equation

$$(A_n+\alpha I)x^{\delta}_{\alpha,n} = z^{\delta}_n$$

for $\alpha \geq \varepsilon_n/c_1$, for $0 < c_1 < 1$.

In order to obtain estimates for $\|\hat{x}-x^{\delta}_{\alpha,n}\|$ we consider the equation

$$(A_n+\alpha I)x_{\alpha,n} = T^*y \ , \quad \alpha \geq \varepsilon_n/c_1 \qquad (7) .$$

Then we have

$$(A_n+\alpha I)(x_{\alpha,n}-x^{\delta}_{\alpha,n}) = T^*y-z^{\delta}_n$$

so that

$$(A+\alpha I)(x_{\alpha,n}-x^{\delta}_{\alpha,n}) = T^*y^\delta - z^\delta_n + (A-A_n)(x_{\alpha,n}-x^{\delta}_{\alpha,n}).$$

Therefore, writing $T^*y-z^\delta_n = T^*(y-y^\delta) + (T^*y^\delta - z^\delta_n)$ and using the relations $\varepsilon_n \leq c_1\alpha$, $0 < c_1 < 1$ and $\|(A+\alpha I)^{-1}\| \leq 1/\alpha$, it follows that

$$\|x_{\alpha,n}-x^{\delta}_{\alpha,n}\| \leq \frac{\delta}{\sqrt{\alpha}} + \frac{\eta_{n,\delta}}{\alpha}.$$

Now, recall equation (2) which can be rewritten as

$$(A_n+\alpha I)x_\alpha = T^*y - (A-A_n)x_\alpha .$$

This together with (7) and using the fact that $\|x_\alpha\| = \|(A+\alpha I)^{-1}T^*y\| = \|(A+\alpha I)^{-1}A\hat{x}\| \leq \|\hat{x}\|$, it follows that

$$\|x_\alpha - x_{\alpha,n}\| \leq \|(A_n+\alpha I)^{-1}\|\|(A-A_n)x_\alpha\| = 0\left[\frac{\varepsilon_n}{\alpha}\right].$$

Thus,

$$\|\hat{x}-x^{\delta}_{\alpha,n}\| = 0\left[\|\hat{x}-x_\alpha\| + \frac{\delta}{\sqrt{\alpha}} + \frac{\eta_{n,\delta}}{\alpha} + \frac{\varepsilon_n}{\alpha}\right] \tag{8}$$

for all $\alpha \geq \varepsilon_n/c_1$.

Now, (8) together with (4), (5) and (6) gives the following result.

THEOREM 1

(i) For $\delta > 0$ and $n=1,2,\ldots$, let $\alpha = \alpha(\delta,n) > 0$ be such that

$$\alpha(\delta,n) \longrightarrow 0 \; , \; \frac{\delta}{\sqrt{\alpha(\delta,n)}} \longrightarrow 0 \; , \; \frac{\varepsilon_n}{\alpha(\delta,n)} \longrightarrow 0 \text{ and } \frac{\eta_{n,\delta}}{\alpha(\delta,n)} \longrightarrow 0$$

as $\delta \longrightarrow 0$, $n \longrightarrow \infty$. Then

$$\|\hat{x}-x^{\delta}_{\alpha,n}\| \longrightarrow 0 \text{ as } \delta \longrightarrow 0 \; , \; n \longrightarrow \infty .$$

(ii) Let $\hat{x} \in R(A^\nu)$, $0 < \nu \leq 1$, and for $\delta > 0$ let $\alpha = c\delta^{2/(2\nu+1)}$ for some constant $c > 0$. Let n_o be a positive integer sucht that for

$$\varepsilon_n \leq c_1\alpha \; , \; \varepsilon_n = 0(\delta^{(2\nu+2)/(2\nu+1)}) \; , \; \eta_{n,\delta} = 0(\delta^{(2\nu+2)/(2\nu+1)}).$$

Then

$$\|\hat{x}-x^{\delta}_{\alpha,n}\| = 0(\delta^{2\nu/(2\nu+1)}) \; , \; n \geq n_o$$

A POSTERIORI PARAMETER CHOICE STRATEGY

For the sake of simplicity of presentations, we assume that

$$\eta_{n,\delta} = 0(\delta^r) \text{ and } \varepsilon_n = 0(\delta^k)$$

for all large enough n, and for some positive real numbers r and k. For $p \geq 0$, $q > 0$, we propose the following "discrepancy principle"

$$\|A_n x^\delta_{\alpha,n} - z^\delta_n\| = \frac{\delta^p}{\alpha^q} \tag{9}$$

where $n > \tilde{n}(\delta)$ with $\tilde{n}(\delta)$ to be specified.

First we note that

$$\|A_n x^\delta_{\alpha,n} - z^\delta_n\| = \|\alpha x^\delta_{\alpha,n}\| = \|\alpha(A_n + \alpha I)^{-1} z^\delta_n\| \leq c \tag{10}$$

for some constant $c > 0$ and for all $\alpha \geq \varepsilon_n/c_1$. Now, let $c_2 > 0$ and $c_3 > 0$ be such that $\|z^\delta_n\| \leq c_2$ for all $n = 1,2,\ldots$ and for all $\delta \leq \delta_o$ for a given $\delta_o > 0$, and $\|A_n\| \leq c_3$ for all $n = 1,2,\ldots$. Then for each $n = 1,2,\ldots$, and for all $\alpha \geq c_4 = \max\left\{\frac{\varepsilon_n}{c_1} : n=1,2,\ldots\right\}$,

$$\|A_n x^\delta_{\alpha,n} - z^\delta_n\| = \|\alpha(A_n + \alpha I)^{-1} z^\delta_n\| \geq \frac{c_2}{1 + c_3/c_4} = c_o \text{, say.} \tag{11}$$

From (10) we have

$$f_n(\alpha) := \alpha^q \|A_n x^\delta_{\alpha,n} - z^\delta_n\| \leq c\alpha^q \text{ , } \alpha \geq \varepsilon_n/c_1.$$

Therefore

$f_n\left[\frac{\varepsilon_n}{c_1}\right] \longrightarrow 0$ as $n \longrightarrow \infty$. Now, let $\tilde{n}(\delta)$ be the smallest positive integer such that for all $n \geq \tilde{n}(\delta)$,

$$\varepsilon_n \leq c_1 \min\left\{\left[\frac{\delta^p}{c}\right]^{1/q}, \left[\frac{\delta^p}{c_o}\right]^{1/q}\right\}.$$

Then taking

$$\alpha_o = \max\left\{c_4, \left[\frac{\delta^p}{c_o}\right]^{1/q}\right\}$$

it follows that for all $n \geq \tilde{n}(\delta)$,

$$\alpha_o \geq \frac{\varepsilon_n}{c_1} \text{ and } f_n\left[\frac{\varepsilon_n}{c_1}\right] \leq \delta^p \leq f_n(\alpha_o).$$

Therefore, by intermediate value theorem, there exists $\alpha = \alpha(\delta,n)$ such that

$$\frac{\varepsilon_n}{c_1} \leq \alpha(\delta,n) \leq \alpha_o \text{ and } \|A_n x^\delta_{\alpha,n} - z^\delta_n\| = \frac{\delta^p}{\alpha^p}$$

for all $n \geq \tilde{n}(\delta)$. We also note that for $n \geq \tilde{n}(\delta)$ and $\alpha = \alpha(\delta,n)$,

$$\begin{aligned}\|z^\delta_n\| - \frac{\delta^p}{\alpha^q} &= \|z^\delta_n\| - \|A_n x^\delta_{\alpha,n} - z^\delta_n\| \\ &\leq \|A_n x^\delta_{\alpha,n}\| \\ &= \frac{1}{\alpha}\|A_n(\alpha x^\delta_{\alpha,n})\| \\ &\leq \|A_n\| \cdot \frac{\delta^p}{\alpha^{q+1}} .\end{aligned}$$

Therefore

$$\alpha^{q+1} \leq \delta^p(\alpha + \|A_n\|)/\|z^\delta_n\| \leq \delta^p\left[\frac{\alpha_o + c_3}{c_2}\right],$$

so that

$$\alpha(\delta,n) = 0(\delta^{p/(q+1)})$$

for all $n \geq \tilde{n}(\delta)$.

We summarize the above results, in Proposition 2 below.

PROPOSITION 2:

Let $p \geq 0$, $q > 0$. For $\delta > 0$, there exists positive integer $\tilde{n}(\delta)$ and $\alpha = \alpha(\delta,n) > 0$ for $n \geq \tilde{n}(\delta)$ such that (9) holds, and

$$\alpha(\delta,n) = 0(\delta^{p/(q+1)}) , \; n \geq \tilde{n}(\delta) .$$

PROPOSITION 3:

If $\frac{p}{q+1} \leq \min\{2,r\}$, then

$$\frac{\delta^p}{\alpha(\delta,n)^q} \leq c\delta^{p/(q+1)}$$

for some constant $c > 0$ independent of n and δ.

Proof:

For any $u \in X$, we have

$$\|(A_n+\alpha I)^{-1}u\| \leq \frac{\|(A+\alpha I)^{-1}u\|}{1-(\varepsilon_n/\alpha)} = \left[\frac{1}{1-c_1}\right]\|(A+\alpha I)^{-1}u\| .$$

Therefore,

$$\|(A_n+\alpha I)^{-1}z_n^\delta\|$$
$$\leq \left[\frac{1}{1-c_1}\right]\|(A+\alpha I)^{-1}(z_n^\delta-T^*y^\delta)+(A+\alpha I)^{-1}T^*(y^\delta-y)+(A+\alpha I)^{-1}T^*y\|$$
$$\leq \left[\frac{1}{1-c_1}\right]\left[\frac{\eta_{n,\delta}}{\alpha} + \frac{\delta}{\sqrt{\alpha}} + \|\hat{x}\|\right],$$

so that

$$\frac{\delta^p}{\alpha^q} = \|\alpha x_{\alpha,n}^\delta\| \leq \frac{1}{1-c_1}(\eta_{n,\delta}+\delta\sqrt{\alpha}+\|\hat{x}\|\alpha) = 0(\delta^\mu)$$
$$\mu = \min\left\{r,\ 1+\frac{p}{2(q+1)},\ \frac{p}{q+1}\right\}.$$

From this the result follows. □

Now we are in a position to state the main theorem of this paper.

THEOREM 4:

If $\frac{p}{q+1} \leq \min\{2,r,k\}$, $\hat{x} \in R(A^\nu)$, $0 < \nu \leq 1$, $\alpha = \alpha(\delta,n)$ is chosen according to (9) for $n \geq \tilde{n}(\delta)$, then

$$\|\hat{x}-x_{\alpha,n}^\delta\| = 0(\delta^s)\ ,\ n \geq \tilde{n}(\delta),$$

where

$$s = \min\left\{\frac{p\nu}{q+1},\ 1-\frac{p}{2(q+1)},\ r-\frac{p}{q+1},\ k-\frac{p}{q+1}\right\}.$$

In particular we have the following:

(a) If $1 + \frac{p}{2(q+1)} \leq r$, $1 + \frac{p}{2(q+1)} \leq k$ and $d\frac{p}{2(q+1)} \leq 1$, then

$$\|\hat{x}-x^{\delta}_{\alpha,n}\| = 0(\delta^{s}),\ s = \begin{cases} \frac{p\nu}{q+1}\ , & \frac{p}{q+1} \leq \frac{2}{2\nu+1} \\ 1 - \frac{p}{2(q+1)}, & \frac{p}{q+1} \geq \frac{2}{2\nu+1} \end{cases} .$$

b) If $r \geq \frac{2\nu+2}{2\nu+1}$, $k \geq \frac{2\nu+2}{2\nu+1}$, then

$$\|\hat{x}-x^{\delta}_{\alpha,n}\| = 0(\delta^{(2\nu/2\nu+1)}) .$$

Proof:

We recall from (5),(6) and (8) that

$$\|\hat{x}-x^{\delta}_{\alpha,n}\| = 0\left[\max\left\{\alpha^{\nu}, \frac{\delta}{\sqrt{\alpha}}, \frac{\eta_{n,\delta}}{\alpha}, \frac{\varepsilon_n}{\alpha}\right\}\right] .$$

From Proposition 3, we have

$$\frac{\delta}{\sqrt{\alpha}} = \delta^{1-p/2(q+1)}\left[\frac{\delta^{p}}{\alpha^{q}}\right]^{1/2q} = 0\left[\delta^{1-p/2(q+1)}\right] ,$$

$$\frac{\delta^{r}}{\alpha} = \delta^{r-p/q}\left[\frac{\delta^{p}}{\alpha^{q}}\right]^{1/q} = 0\left[\delta^{r-p/(q+1)}\right]$$

and

$$\frac{\delta^{k}}{\alpha} = 0\left[\delta^{k-p/(q+1)}\right] .$$

These estimates together with Proposition 2 and the assumptions $\eta_{n,\delta} = 0(\delta^{r})$ and $\varepsilon_n = 0(\delta^{k})$ imply the results.

□

It is to be remarked that in the projection method case $z_n^\delta = P_n T^* y^\delta$ and in the degenerate kernel method, $z_n^\delta = T^* y^\delta$, so that in both the above cases,

$$\|z_n^\delta - T^* y^\delta\| \longrightarrow 0 \ , \quad \text{as} \quad n \longrightarrow \infty \, .$$

Moreover, in both the above cases the operator A_n is non–negative and self–adjoint with $\|A - A_n\| \longrightarrow 0$, so that the analyis can be simplified.

Thus this paper generalizes the type of results in Engl and Neubauer ([2],[3]) for projection methods, and also the first part of Grotsch [8] providing a parameter choice strategy.

ACKNOWLEDGEMENT

This work is carried out while Thamban Nair was a Visiting Professor at the Universität Kaiserslautern, Fachbereich Mathematik. The support received is gratefully acknowledged.

REFERENCES

[1] Engl, H.W.: Discrepancy principles for Tikhonov regularization of ill–posed problems leading to optimal convergence rates. J. Optim. Th. and Appl. 52 (1987), 209–215.

[2] Engl, H.W. and A. Neubauer: An improved version of Marti's method for solving ill–posed linear integral equations. Math. Computation, 45 (1985), 405–416.

[3] Engl, H.W. and A. Neubauer: Optimal parameter choice for ordinary and iterated Tikhonov regularization. In: Inverse and Ill–Posed Problems (Eds.: H.W. Engl and C.W. Groetsch) Academic Press, Inc. London, 1987, 97–125.

[4] Gfrerer, H.: Parameter choice for Tikhonov regularization of ill–posed problems. In: Inverse and ill–posed problems (Eds.: H.W. Engl and C.W. Groetsch) Academic Press, Inc. London, 1987, 127–149.

[5] Groetsch, C.W.: Comments on Morozov's discrepancy principle. In: Improperly Posed Problems and Their Numerical Treatment. (Eds.: G. Hämmerlin and K.H. Hoffmann), Birkhäuser, 1983, 97–104.

[6] Groetsch, C.W.: The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind. Pitman Books, London, 1984.

[7] Groetsch, C.W.: Convergence analysis of a regularized degenerate kernel method for Fredholm integral equations of the first kind. Integral Equations and Operator Theory 13 (1990), 67–75.

[8] Groetsch, C.W. and J. Guacaneme: Regularized Ritz approximations for Fredholm equations of the first kind. Rocky Mountain J. Math. 15,1 (1985), 33–37.

[9] Nair, M.T.: A generalization of Arcangeli's method for ill–posed problems leading to optimal rates. Integral Equations and Operator Theory (To appear).

[10] Nair, M.T.: A unified approach for regularized approximation methods for Fredholm integral equations of the first kind (Communicated).

[11] Schock, E.: On the asymptotic order of accuracy of Tikhonov regularizations. J. Optim. Th. and Appl. 44 (1984), 95–104.

[12] Schock, E.: Parameter choice by discrepancy principle for the approximate solution of ill–posed problems. Integral Equations and Operator Theory 7 (1984), 895–898.

[13] Schock, E.: What are the proper condition numbers of discretized ill–posed problems ? Lin. Alg. and Appl. 81 (1986), 129–136.