A discrepancy principle for Tikhonov regularization
with approximately specified data

by M. Thamban Nair (Chennai) and
Eberhard Schock (Kaiserslautern)

Abstract. Many discrepancy principles are known for choosing the parameter \( \alpha \) in the regularized operator equation \((T^*T+\alpha I)x_\alpha = T^*y_\delta, \|y-y_\delta\| \leq \delta\) in order to approximate the minimal norm least-squares solution of the operator equation \(Tx = y\). We consider a class of discrepancy principles for choosing the regularization parameter when \(T^*T\) and \(T^*y_\delta\) are approximated by \(A_n\) and \(z_n^\delta\) respectively with \(A_n\) not necessarily self-adjoint. This procedure generalizes the work of Engl and Neubauer (1985), and particular cases of the results are applicable to the regularized projection method as well as to a degenerate kernel method considered by Groetsch (1990).

1. Introduction. We are concerned with the problem of finding approximations to the minimal norm least-squares solution \(\hat{x}\) of the operator equation

\[Tx = y,\]

where \(T : X \to Y\) is a bounded linear operator between Hilbert spaces \(X\) and \(Y\), and \(y\) belongs to \(D(T^\dagger) := R(T)+R(T)^\perp\), the domain of the Moore–Penrose inverse \(T^\dagger\) of \(T\). It is well known [8] that if the range \(R(T)\) of \(T\) is not closed, then the operator \(T^\dagger\) which associates \(y \in D(T^\dagger)\) to \(\hat{x} := T^\dagger y\), the unique least-squares solution of minimal norm, is not continuous, and consequently the problem of solving (1.1) for \(\hat{x}\) is ill-posed. A prototype of an ill-posed problem is the Fredholm integral equation of the first kind

\[
\int_0^1 k(s,t)x(t)\ dt = y(s), \quad 0 \leq s \leq 1,
\]

1991 Mathematics Subject Classification: 65J10, 65R30, 45B05, 45E99.
Key words and phrases: ill-posed problems, minimal norm least-squares solution, Moore–Penrose inverse, Tikhonov regularization, discrepancy principle, optimal rate.

The work of M. Thamban Nair is partially supported by a project grant from National Board for Higher Mathematics, Department of Atomic Energy, Govt. of India.
with nondegenerate kernel \(k(\cdot, \cdot) \in L^2([0,1] \times [0,1])\), where \(X = Y = L^2[0,1]\). Regularization methods are employed to find approximations to \(\hat{x}\). In Tikhonov regularization one looks for the unique \(x_\alpha, \alpha > 0\), which minimizes the functional
\[
x \to ||Tx - y||^2 + \alpha||x||^2, \quad x \in X.
\]
Equivalently, one solves the well-posed equation
\[
(T^*T + \alpha I)x_\alpha = T^*y
\]
for each \(\alpha > 0\). Since \(T^*T\hat{x} = T^*y\), it follows that
\[
||\hat{x} - x_\alpha|| = ||\alpha(T^*T + \alpha I)^{-1}\hat{x}|| \leq ||\hat{x}||.
\]
It is known ([8], [16]) that
\[
\alpha \rightarrow ||\hat{x} - x_\alpha|| \rightarrow 0 \quad \text{as} \quad \alpha \rightarrow 0
\]
and
\[
\hat{x} \in R((T^*T)^\nu), \quad 0 \leq \nu \leq 1, \quad \text{implies} \quad ||\hat{x} - x_\alpha|| = O(\alpha^\nu).
\]
In practical applications the data \(y\) may not be available exactly, instead one may have an approximation \(y^\delta\) with say \(||y - y^\delta|| \leq \delta, \delta > 0\). Then one solves the equation
\[
(T^*T + \alpha I)x_\alpha^\delta = T^*y^\delta
\]
instead of (1.3) and requires \(||\hat{x} - x_\alpha^\delta|| \rightarrow 0 \quad \text{as} \quad \alpha \rightarrow 0 \quad \text{and} \quad \delta \rightarrow 0\). It follows from (1.3) and (1.7) that
\[
||x_\alpha - x_\alpha^\delta||^2 = ||(T^*T + \alpha I)^{-1}T^*(y - y^\delta)||^2
\]
\[
= ||(T^*T + \alpha I)^{-1}T^*(y - y^\delta)||^2
\]
\[
= ||(TT^* + \alpha I)^{-2}TT^*(y - y^\delta)||^2
\]
\[
\leq ||(TT^* + \alpha I)^{-2}TT^*|| \cdot ||(y - y^\delta)||^2 \leq \delta^2/\alpha,
\]
so that
\[
||\hat{x} - x_\alpha^\delta|| \leq ||\hat{x} - x_\alpha|| + \delta/\sqrt{\alpha}.
\]
Now let \(R_\alpha = (T^*T + \alpha I)^{-1}T^*\) for \(\alpha > 0\). Then by (1.5) we have
\[
||R_\alpha y - T^*y|| \rightarrow 0 \quad \text{as} \quad \alpha \rightarrow 0
\]
for \(y \in D(T^*)\). Therefore, if \(R(T)\) is not closed, then the family \(\{R_\alpha\}_{\alpha > 0}\) is not uniformly bounded so that, as a consequence of the Uniform Boundedness Principle, there exists \(v \in Y\) such that \(\{R_\alpha v\}_{\alpha > 0}\) is not bounded in \(Y\). In particular, if \(y^\delta = y + \delta v/\|v\|\), then \(||y - y^\delta|| \leq \delta\) and \(\{R_\alpha y^\delta\}_{\alpha > 0}\) is unbounded in \(Y\). Therefore, the problem of choosing the regularization parameter \(\alpha\) depending on \(y^\delta\) is important. Many works in the literature are devoted to this aspect (cf. [7], [17], [1], [2], [3], [6], [14], [4]).
In order to solve (1.7) numerically, it is required to consider approximations of $T^*T$ and of $T^*y^\delta$. So the problem actually at hand would be of the form

$$ (A_n + \alpha I)x_{\alpha,n}^\delta = z_n^\delta, $$

where $(A_n)$ and $(z_n^\delta)$ are approximations of $T^*T$ and of $T^*y^\delta$ respectively.

In the well known regularized projection methods (cf. [10], [2], [3]),

$$ A_n = P_nT^*TP_n \quad \text{and} \quad z_n^\delta = P_nT^*y^\delta, $$

where $(P_n)$ is a sequence of orthogonal projections on $X$ such that $P_n \to I$ pointwise. In this case we have

$$ \|T^*T - A_n\| \to 0 \quad \text{as} \quad n \to \infty, $$

and discrepancy principles are known for choosing the regularization parameter $\alpha$ in (1.9) (see e.g. [2], [3], [13], [5]).

In the degenerate kernel methods for the integral equation (1.2) with $k(\cdot, \cdot) \in C([0,1] \times [0,1])$, $A_n$ is obtained by approximating the kernel $\tilde{k}(\cdot, \cdot)$ of the integral operator $T^*T$ by a degenerate kernel $\tilde{k}_n(\cdot, \cdot)$ so that $\|\tilde{k} - \tilde{k}_n\|_{\infty} \to 0$ as $n \to \infty$. Then it follows that

$$ \|T^*T - A_n\| \leq \|\tilde{k} - \tilde{k}_n\|_2 \leq \|\tilde{k} - \tilde{k}_n\|_{\infty} \to 0 \quad \text{as} \quad n \to \infty. $$

(See [11] and [12] for a discussion on degenerate kernel methods for integral equations.) In a degenerate kernel method considered by Groetsch [9] the approximation $\tilde{k}_n(\cdot, \cdot)$ is obtained from

$$ \tilde{k}(s, t) := \int_0^1 k(\tau, s)k(\tau, t) \, d\tau, \quad a \leq s, t \leq b. $$

by using a convergent quadrature rule. In this case one has $\|\tilde{k} - \tilde{k}_n\|_{\infty} \to 0$ as $n \to \infty$ for nice enough kernels $k(\cdot, \cdot)$.

Moreover, for the degenerate kernel method of Groetsch [9] as well as for the regularized projection methods, the operators $A_n$ are non-negative and self-adjoint.

In this paper we consider the generalized form of a class of discrepancy principles in [1], namely,

$$ \|A_n x_{\alpha,n}^\delta - z_n^\delta\| = \frac{\delta^p}{\alpha^q}, \quad p > 0, \quad q > 0, $$

for large enough $n$, to choose the regularization parameter $\alpha = \alpha(n, \delta)$ in (1.9), where $(A_n)$ is a sequence of bounded linear operators on $X$ and $(z_n^\delta)$ in $X$ such that

$$ \|T^*T - A_n\| \to 0 \quad \text{and} \quad \|T^*y^\delta - z_n^\delta\| \to 0 \quad \text{as} \quad n \to \infty. $$
It has to be observed that we do not assume the operators $A_n$ to be non-negative and self-adjoint. The consideration of a general $A_n$, as has been done recently by Nair [15], is important from the computational point of view, because even if one starts with a non-negative self-adjoint operator as approximation of $T^*T$, due to truncation errors etc., one actually may not be dealing with a non-negative self-adjoint operator.

With $\alpha$ chosen according to (1.10), we show the convergence of the solution $x^{\alpha,n}_\delta$ of (1.9) to $\hat{x}$ as $\delta \to 0$, $n \to \infty$, and also obtain estimates for the error $\|\hat{x} - x^{\alpha,n}_\delta\|$ whenever $\hat{x} \in R((T^*T)^\nu)$, $0 < \nu \leq 1$. Our result on error estimates shows that if $\nu_0$ is an estimate for the possibly unknown $\nu$, with $0 < \nu \leq \nu_0 \leq 1$, then taking $p/(q + 1) = 2/(2\nu_0 + 1)$ one obtains the rate $O(\delta^{2
u/(2\nu+1)})$. In particular, prior knowledge of $\nu$ enables us to obtain the optimal rate $O(\delta^{2
u/(2\nu+1)})$ (cf. Schock [16]).

If $A_n = P_n T^* T P_n$ and $z^{\delta}_n = P_n T^* y^{\delta}$ then (1.10) coincides with a discrepancy principle considered by Engl and Neubauer [2] and we recover the optimal result in [2] as a particular case. Thus this paper generalizes the type of results in [2] and [9] for projection methods and degenerate kernel method for integral equations respectively, providing also a parameter choice strategy in the latter case.

2. Approximate solution and convergence. Let $X$ and $Y$ be Hilbert spaces and $T : X \to Y$ be a bounded linear operator with its range $R(T)$ not necessarily closed in $Y$. Let $y \in D(T^\dagger) := R(T) + R(T)^\perp$, $y \neq 0$, so that there exists a unique $\hat{x} \in X$ of minimal norm such that

$$\|Tx - y\| = \inf \{\|Tx - y\| : x \in X\}.$$  

Let $(A_n)$ be a sequence of bounded linear operators on $X$ and for $\delta > 0$, let $y^{\delta} \in Y$ and $(z^{\delta}_n)$ in $X$ be such that

$$\|T^* T - A_n\| \leq \varepsilon_n, \quad \|y - y^{\delta}\| \leq \delta, \quad \|T^* y^{\delta} - z^{\delta}_n\| \leq \eta^{\delta}_n,$$

where $(\varepsilon_n)$ and $(\eta^{\delta}_n)$ are sequences of nonnegative real numbers such that

$$\varepsilon_n \to 0 \quad \text{as} \quad n \to \infty$$

and

$$\eta^{\delta}_n \to 0 \quad \text{as} \quad n \to \infty \text{ and } \delta \to 0.$$  

Throughout the paper we denote the operator $T^* T$ by $A$, and $c$, $c'$, $c_1$, $c_2$, etc., denote positive constants which may assume different values in different contexts.

**THEOREM 2.1.** If $\varepsilon_n \leq c_0 \alpha$ with $0 < c_0 < 1$, then $A_n + \alpha I$ is bijective and

$$\|(A_n + \alpha I)^{-1}\| \leq 1/(\alpha(1 - c_0)).$$
Moreover, if \( x_\alpha^\delta \) and \( x_{\alpha,n}^\delta \) are the unique solutions of (1.7) and (1.9) respectively, then

\[
\| \hat{x} - x_{\alpha,n}^\delta \| \leq c\left( \| \hat{x} - x_\alpha^\delta \| + \frac{\eta_n}{\alpha} + \frac{\varepsilon_n}{\alpha} \right).
\]

In particular, if \( \alpha := \alpha(\delta, n) \) is chosen in such a way that

\[
\alpha(\delta, n) \to 0, \quad \frac{\delta}{\sqrt{\alpha(\delta, n)}} \to 0, \quad \frac{\varepsilon_n}{\alpha(\delta, n)} \to 0 \quad \text{and} \quad \frac{\eta_n}{\alpha(\delta, n)} \to 0
\]

as \( \delta \to 0 \) and \( n \to \infty \), then

\[
\| \hat{x} - x_{\alpha,n}^\delta \| \to 0 \quad \text{as} \quad \delta \to 0 \quad \text{and} \quad n \to \infty.
\]

\textbf{Proof.} Since \( A \) is non-negative and self-adjoint, it follows from spectral theory that for each \( \alpha > 0 \), \( (A + \alpha I)^{-1} \) exists as a bounded linear operator on \( X \) and

\[
\|(A + \alpha I)^{-1}\| \leq 1/\alpha.
\]

Therefore, if \( \|A - A_n\| < 1/\|(A + \alpha I)^{-1}\| \) then, by results on perturbation of operators, \( (A_n + \alpha I)^{-1} \) exists and is a bounded operator, and

\[
\|(A_n + \alpha I)^{-1}\| \leq \frac{\|(A + \alpha I)^{-1}\|}{1 - \|A - A_n\| \cdot \|(A + \alpha I)^{-1}\|} \leq \frac{1/\alpha}{1 - \varepsilon_n/\alpha} \leq \frac{1}{\alpha(1 - c_0)}.
\]

Now let \( w_{\alpha,n}^\delta \) be the unique solution of the equation (1.9) with \( T^* y^\delta \) in place of \( z_n^\delta \), i.e.,

\[
(2.3) \quad (A_n + \alpha I)w_{\alpha,n}^\delta = T^* y^\delta.
\]

Then from (1.7), (1.9) and (2.3), we have

\[
x_{\alpha,n}^\delta - w_{\alpha,n}^\delta = (A_n + \alpha I)^{-1}(z_n^\delta - T^* y^\delta)
\]

and

\[
w_{\alpha,n}^\delta - x_\alpha^\delta = (A_n + \alpha I)^{-1}(A - A_n)x_\alpha^\delta.
\]

Since \( \varepsilon_n \leq c_0\alpha \), it follows that

\[
\|x_{\alpha,n}^\delta - w_{\alpha,n}^\delta\| \leq c_1\eta_n/\alpha
\]

and

\[
\|w_{\alpha,n}^\delta - x_\alpha^\delta\| \leq c_2(\|\hat{x} - x_\alpha^\delta\| + \varepsilon_n/\alpha),
\]

so that

\[
\|\hat{x} - x_{\alpha,n}^\delta\| \leq c(\|\hat{x} - x_\alpha^\delta\| + \eta_n/\alpha + \varepsilon_n/\alpha).
\]

Now the assumptions on \( \alpha := \alpha(\delta, n) \) together with (1.6) and (1.8) imply the convergence \( \|\hat{x} - x_{\alpha,n}^\delta\| \to 0 \) as \( \delta \to 0 \) and \( n \to \infty \).
3. The discrepancy principle. By our assumption (2.1) on \( \eta_n^\delta \) and the fact that \( 0 \neq y \in D(T^\dagger) \), we have \( c_1 \leq \|z_n^\delta\| \leq c_2 \) for all large enough \( n \), say \( n \geq n_0(\delta) \) and for each \( \delta \in (0, \delta_0) \) for some \( \delta_0 \). Therefore by Theorem 2.1, 
\[
\|A_n x_{\alpha,n}^\delta - z_n^\delta\| = \|\alpha x_{\alpha,n}^\delta\| = \|\alpha(A_n + \alpha I)^{-1} z_n^\delta\| \leq \gamma_1 
\]
for some constant \( \gamma_1 \) and for all \( \alpha \geq \varepsilon_n/c_0 \). Moreover, if 
\[
\alpha \geq \gamma_0 := \max\{\varepsilon_n/c_0 : n = 1, 2, \ldots\} \quad \text{and} \quad \delta \leq \delta_0,
\]
then 
\[
\|A_n x_{\alpha,n}^\delta - z_n^\delta\| \geq \|\alpha(A_n + \alpha I)^{-1} z_n^\delta\| \geq \frac{\gamma_0 \|z_n^\delta\|}{\|A_n\| + \alpha} \geq \gamma_2 
\]
for some \( \gamma_2 > 0 \), since \( (A_n) \) is uniformly bounded.

Now to choose the regularization parameter \( \alpha \) in (1.9), we consider the discrepancy principle (1.10).

For simplicity of presentation we assume that 
\[
\eta_n^\delta \leq c_3 \delta^r \quad \text{and} \quad \varepsilon_n \leq c_4 \delta^k 
\]
for some positive reals \( r \) and \( k \), and for all \( n \geq n_0(\delta) \).

**Theorem 3.1.** Let \( p \) and \( q \) be positive integers. Then for each \( \delta \in (0, \delta_0) \), there exists a positive integer \( n_1(\delta) \) and for each \( n \geq n_1(\delta) \), there exists \( \alpha := \alpha(\delta, n) \) such that (1.10) is satisfied. Moreover, 
\[
\alpha \leq c_1 \delta^{p/(q+1)} \quad \text{and} \quad \delta^p/\alpha^q \leq c_2 \delta^\mu, \quad n \geq n_1(\delta),
\]
where 
\[
\mu = \min \left\{ r, \frac{p}{(q+1)}, 1 + \frac{p}{2(q+1)} \right\}.
\]

**Proof.** Let \( \delta \in (0, \delta_0) \). For \( \alpha \geq \varepsilon_n/c_0 \) and \( n = 1, 2, \ldots \), define 
\[
f_n(\alpha) = \alpha^q \|A_n x_{\alpha,n}^\delta - z_n^\delta\|.
\]
Then from (3.1) it follows that 
\[
f_n(\varepsilon_n/c_0) \to 0 \quad \text{as} \quad n \to \infty.
\]
Let \( n_1(\delta) \geq n_0(\delta) \) be the smallest positive integer such that for all \( n \geq n_1(\delta) \), 
\[
\varepsilon_n \leq c_0 \min\{((\delta^p/\gamma_2)^{1/q}, (\delta^p/\gamma_1)^{1/q}\}
\]
Then taking \( \alpha_0 = \max\{\gamma_0,(\delta^p/\gamma_2)^{1/q}\} \), we obtain 
\[
\varepsilon_n \leq c_0 \alpha_0 \quad \text{and} \quad \alpha_0 \geq \gamma_0
\]
so that by (3.1) and (3.2), we have 
\[
f_n(\varepsilon_n/c_0) \leq \delta^p \leq f_n(\alpha_0).
\]
Therefore by the Intermediate Value Theorem, there exists \( \alpha := \alpha(\delta, n) \) such that 
\[
\varepsilon_n/c_0 \leq \alpha \leq \alpha_0 \quad \text{and} \quad \|A_n x_{\alpha,n}^\delta - z_n^\delta\| = \delta^p/\alpha^q
\]
A discrepancy principle for Tikhonov regularization

for all \( n \geq n_1(\delta) \). We also note that

\[
x^\delta_{\alpha,n} = \frac{1}{\alpha}(z^\delta_n - A_n x^\delta_{\alpha,n})
\]

so that for all \( n \geq n_1(\delta) \) and \( \alpha = \alpha(\delta,n) \),

\[
\|z^\delta_n\| - \delta^p/\alpha^q = \|z^\delta_n\| - \|A_n x^\delta_{\alpha,n} - z^\delta_n\| \leq \|A_n x^\delta_{\alpha,n}\| \leq \|A_n\|\delta^p/\alpha^q + 1.
\]

Therefore \( \alpha^q + 1 \leq \delta^p(\alpha + \|A_n\|)/\|z^\delta_n\| \leq c\delta^p \) and consequently

\[
\alpha(\delta,n) \leq c_1\delta^p/(q+1), \quad n \geq n_1(\delta).
\]

Now, using the estimates in (1.4), (1.8) and (2.2), we have

\[
\frac{p}{q+1} \leq \min\{2,r,k\},
\]

where \( r \) and \( k \) are as in (3.3).

**Theorem 4.1** Let \( \alpha := \alpha(\delta,n) \) be chosen according to (1.10). Then:

(i) \( \|\hat{x} - x^\delta_{\alpha,n}\| \to 0 \) as \( n \to \infty \) and \( \delta \to 0 \).

(ii) If \( \hat{x} \in R(A^\nu) \), \( 0 < \nu \leq 1 \), then for all large enough \( n \) and small enough \( \delta \),

\[
\|\hat{x} - x^\delta_{\alpha,n}\| \leq c\delta^s,
\]

where

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

(iii) In particular, if

\[
\min\{r,k\} \geq \frac{2
\nu + 2}{2\nu+1} \quad \text{and} \quad \frac{p}{q+1} = \frac{2}{2\nu+1},
\]

then

\[
\|\hat{x} - x^\delta_{\alpha,n}\| \leq c\delta^{2\nu/(2\nu+1)}.
\]
Proof. Using (3.4), we have
\[ \frac{\delta^l}{\alpha^m} = \frac{\delta^{l-mp/q}(\delta p/\alpha q)^m}{q} \leq c\delta^{l-m(p-\mu)/q} \]
for every \( l \geq 0 \) and \( m \geq 0 \), where \( \mu \) is as in Theorem 3.1. But by the assumption (4.1), \( \mu = \frac{p}{q+1} \), so that
\[ \frac{\delta^l}{\alpha^m} \leq c\delta^{l-mp/(q+1)}. \]
Therefore
\[ \frac{\delta}{\sqrt{\alpha}} \leq c_1\delta^{1-p/2(q+1)}, \quad \frac{\eta_n^\delta}{\alpha} \leq c_2\delta^{r-p/(q+1)} \quad \text{and} \quad \frac{\varepsilon_n}{\alpha} \leq c_3\delta^{k-p/(q+1)}. \]
Using this, the result in (i) follows from (1.5), (1.8) and (2.2), the estimate in (ii) follows from (1.6), (1.8) and (2.2), and (iii) is a consequence of (ii).

Acknowledgements. The first version of this paper was written while M. Thamban Nair was a Visiting Professor at the Fachbereich Mathematik, Universität Kaiserslautern, Germany. The support received is gratefully acknowledged.

References

A discrepancy principle for Tikhonov regularization


Department of Mathematics
Indian Institute of Technology Madras
Chennai 600 036, India
E-mail: mtnair@acer.iitm.ernet.in

Fachbereich Mathematik
Universität Kaiserslautern
Kaiserslautern, Germany
E-mail: schock@mathematik.uni-kl.de

Reçu par la Rédaction le 21.8.1995
Révisé le 10.5.1998