Monograph:

Dodd, T.J. and Harrison, R.F. (2002) Steepest descent for a linear operator equation of the second kind with application to Tikhonov regularisation. Research Report. ACSE Research Report no. 827 . Automatic Control and Systems Engineering, University of Sheffield

Steepest Descent for a Linear Operator Equation of the Second Kind with Application to Tikhonov Regularization

Tony J. Dodd and Robert F. Harrison
Department of Automatic Control and Systems Engineering
The University of Sheffield, Sheffield S1 3JD, UK
e-mail: {t.j.dodd, r.f.harrison}@shef.ac.uk

Research Report No. 827
October 2002

Abstract

Let H_1, H_2 be Hilbert spaces, T a bounded linear operator on H_1 into H_2 such that the range of T, $\mathcal{R}(T)$, is closed. Let T^* denote the adjoint of T. In this paper, we review the generalised solution, and method of steepest descent, for the linear operator equation, $Tx = b, b \in H_2$. Further, we establish the convergence of the method of steepest descent to the unique solution $(T^*T + \lambda I)^{-1}T^*b, b \in H_2$ of the operator equation of the second kind, $(T^*T + \lambda I)x = T^*b$, if x_0 is in the range of T^*. This new result is shown to have immediate application to the iterative solution of the Tikhonov regularisation method for the original operator equation, $Tx = b, b \in H_2$.

1 Introduction

Following the discussions in (Kammerer and Nashed 1972; Nashed 1970), let H_1 and H_2 be Hilbert spaces over the same scalars (real or complex). For any subspace, S, of H_1 or H_2, the orthogonal complement and closure of S are denoted by $S^\perp$ and $\bar{S}$ respectively. We consider a bounded linear operator, T, on H_1 into H_2. Then T^* denotes the adjoint of T, i.e. for all $x \in H_1, b \in H_2$ [1],

$$\langle Tx, b\rangle = \langle x, T^*b\rangle.$$

Let $\mathcal{R}(T)$ and $\mathcal{N}(T)$ denote, respectively, the range and null spaces of T. The following relations are then well known (Nashed 1970)

$$H_1 = \mathcal{N}(T) \oplus \mathcal{N}(T)^\perp, \tag{1}$$
$$H_2 = \mathcal{N}(T^*) \oplus \mathcal{N}(T^*)^\perp, \tag{2}$$
$$\{\overline{\mathcal{R}(T)}\}^\perp = \mathcal{N}(T^*), \qquad \overline{\mathcal{R}(T^*)} = \mathcal{N}(T)^\perp, \tag{3}$$
$$\mathcal{R}(T) \text{ is closed} \quad\Leftrightarrow\quad \mathcal{R}(T^*) \text{ is closed ,} \tag{4}$$
$$\mathcal{N}(T^*T) = \mathcal{N}(T), \qquad \overline{\mathcal{R}(T)} = \overline{\mathcal{R}(TT^*)}. \tag{5}$$

Definition 1.1 *A vector $u \in H_1$ is called a least squares solution of the linear operator equation*

$$Tx = b, \quad b \in H_2, \tag{6}$$

if $\inf\{\|Tx - b\| : x \in H_1\} = \|Tu - b\|$.

It can be shown that u is a least squares solution of (6) if, and only if, u is a solution of the "normal" equation

$$T^*Tx = T^*b. \tag{7}$$

The following theorem summarises this.

Theorem 1.1 (Groetsch, 1977) *Suppose $T : H_1 \to H_2$ has closed range and $b \in H_2$, then the following conditions on $x \in H_1$ are equivalent:*

(i) $Tx = Pb$;

(ii) $\|Tu - b\| \le \|Tx - b\|$ *for any* $x \in H_1$; *and*

(iii) $T^*Tx = T^*b$.

where P denotes the projection of b onto $\mathcal{N}(T^*)^\perp = \mathcal{R}(T)$.

In the case that H_1 is finite dimensional (7) always has at least one solution since $\mathcal{R}(T^*) = \mathcal{R}(T^*T)$. However, when H_1 is infinite dimensional and the range, $\mathcal{R}(T)$, is not closed this equation may have no solutions. We therefore assume that $\mathcal{R}(T)$ is closed. Then there always exists at least one least squares

[1] In expressing inner products, $\langle\cdot,\cdot\rangle$, it is assumed the Hilbert space to which the inner product belongs is obvious.

solution for each $b \in H_2$. For $\mathcal{N}(T) \neq \{0\}$ there will be infinitely many solutions since if u is a least squares solution so is $u + w$ for any $w \in \mathcal{N}(T)$.

By the continuity and linearity of T and T^*, the set, S, of all least squares solutions of (6) is a nonempty closed, convex set. Hence it contains a unique element, v, of minimal norm, i.e.,

$$||Tv - b|| \leq ||Tx - b|| \quad \text{for all } x \in H_1,$$

and

$$||v|| < ||u|| \quad \text{for all } u \in S, \quad u \neq v.$$

We then define the generalised inverse of T as the operator $T^\dagger : H_2 \to H_1$ such that $T^\dagger b = v$, i.e. that operator which assigns, to each $b \in H_2$, the unique least squares solution of minimal norm of (6). $T^\dagger$ is linear and bounded. Note that if T is invertible we have $T^\dagger = T^{-1}$. The associated least squares solution of minimal norm is the generalised solution, denoted $x^\dagger (= v)$. It can also be shown that $T^\dagger b$ is the unique least squares solution in $\mathcal{R}(T^*)$ (Nashed 1970).

We will need the following lemma in Section 4.

Lemma 1.1 (Petryshyn, 1967) *Suppose T is as described above and $\mathcal{R}(T)$ is closed. Then the restriction of T to $N(T)^\perp = \mathcal{R}(T^*)$ has a bounded inverse. Equivalently, there exists a number $m > 0$ such that $||Tx|| \geq m||x||$ for all $x \in N(T)^\perp = \mathcal{R}(T^*)$.*

For a detailed introduction to the theory of generalised inverses of linear operators with closed and arbitrary range, see (Groetsch 1977). Additional useful references include (Ben-Israel and Charnes 1963) which describes, specifically, the Euclidean case, (Desoer and Whalen 1963; Nashed 1971) and the volume (Nashed 1976). In the next section a particular class of linear operator equations, those of the second kind, is described. Under certain, shown, conditions such equations are uniquely invertible and therefore the generalised inverse is the true inverse. The method of steepest descent for the iterative solution of linear operator equations is summarised in Section 3 (Nashed 1970; McCormick and Rodrigue 1975). For the class of equations of the second kind considered an alternative method of steepest descent is presented in Section 4. It is shown, using an extension of the proof in (Nashed 1970), that the steepest descent iteration converges to the unique solution. In Section 5 the operator equation of the second kind is applied to the Tikhonov regularisation of the original linear operator equation. A natural corollary of the new steepest descent method is therefore an iterative solution to the Tikhonov regularisation method.

2 Operator Equations of the Second Kind

Consider now the linear operator equation

$$Ax - \lambda x = y \tag{8}$$

where $x, y \in H_1$, A is a linear operator on H_1 into H_1, λ is a (complex) number and I is the identity operator on H_1. This equation is called an operator equation of the second kind, by analogy with integral equations (in contrast we call (6) an operator equation of the first kind) (Kammerer and Nashed 1972; Kantorovich and Akilov 1964).

We assume now that the operator $A = T^*T$ where T is as in Section 1, i.e. linear and bounded. A is then itself bounded and also self-adjoint and nonnegative (often simply referred to as positive), the latter since

$$\langle Ax, x\rangle = \langle T^*Tx, x\rangle = \langle Tx, Tx\rangle = \|Tx\| \geq 0. \tag{9}$$

Theorem 2.1 *Assume that $\lambda > 0$, then the equation*

$$(T^*T + \lambda I)x = y$$

is uniquely solvable for all $y \in H_1$, i.e. the inverse $(T^*T + \lambda I)^{-1}$ always exists.

Proof. We require that $T^*T + \lambda I$ is one-to-one for every $\lambda > 0$. Let $T^*Tx + \lambda x = 0$. Multiplication by x gives $\langle T^*Tx, x\rangle + \langle \lambda x, x\rangle = 0$. By the definition of the adjoint $\langle Tx, Tx\rangle + \lambda\langle x, x\rangle = 0$ and therefore we must have $x = 0$. □

Suppose, further, that $y = T^*b, b \in H_2$, then we have

Corollary 2.1 *For every $b \in H_2$ and $\lambda > 0$, the equation*

$$(T^*T + \lambda I)x = T^*b$$

is uniquely solvable. The solution can be written in the form $x = (T^*T + \lambda I)^{-1}T^*b$.

Theorem 2.2 *For every $b \in H_2$ and $\lambda > 0$, the solution of the equation*

$$(T^*T + \lambda I)x = T^*b$$

can be written, equivalently, in the form $x = T^*(TT^* + \lambda I)^{-1}b$ and therefore we must have $x \in R(T^*)$.

Proof. Consider the operator $T^*TT^* + \lambda T^*$, which gives rise to the following equation

$$T^*(TT^* + \lambda I) = (T^*T + \lambda I)T^*$$

where it is assumed that the appropriate identity operator, I, is used on each side. Premultiplying by $(T^*T + \lambda I)^{-1}$,

$$(T^*T + \lambda I)^{-1}T^*(TT^* + \lambda I) = T^*$$

and subsequently postmultiplying by $(TT^* + \lambda I)^{-1}$,

$$(T^*T + \lambda I)^{-1}T^* = T^*(TT^* + \lambda I)^{-1}.$$

But we already have $x = (T^*T + \lambda I)^{-1}T^*b$ (Corollary 2.1) and therefore the unique solution must also be expressible as $x = T^*(TT^* + \lambda I)^{-1}b$. It is then obvious that $x \in R(T^*)$. □

3 The Method of Steepest Descent

Let T be a bounded linear operator on H_1 into H_2 and assume that $\mathcal{R}(T)$ is closed. Suppose that $J : H_1 \to \mathbb{R}$ is the non-negative functional

$$J(x) = \frac{1}{2}\|Tx - b\|^2. \tag{10}$$

We seek a point $x^* \in H_1$ such that

$$J(x^*) = \inf\{J(x) : x \in H_1\}.$$

The method of steepest descent for minimising the functional, $J(x)$, for $b \in H_2$ is defined, for a given initial approximation, x_0, by the following sequence (Nashed 1970; McCormick and Rodrigue 1975)

$$x_{n+1} = x_n - \gamma_n R_n, \quad n = 0, 1, \ldots \tag{11}$$

where

$$R_n = T^*Tx_n - T^*b = T^*(Tx_n - b) \tag{12}$$

and

$$\gamma_n = \frac{\|R_n\|^2}{\|TR_n\|^2}. \tag{13}$$

Theorem 3.1 (Nashed, 1970) *Let H_1 and H_2 be Hilbert spaces and T be a bounded linear operator on H_1 into H_2 such that its range, $\mathcal{R}(T)$, is closed. The sequence of steepest descent defined by (11)-(13) converges to a least squares solution of $Tx = b$ for any $x_0 \in H_1$. The sequence $\{x_n\}$ converges to the unique element $T^\dagger b$ if, and only if, $x_0 \in \mathcal{R}(T^*)$.*

4 Steepest Descent for Equations of the Second Kind

Again, let T be a bounded linear operator on H_1 into H_2 and assume that $\mathcal{R}(T)$ is closed. Suppose now that instead of $J(x)$ we have the non-negative functional

$$J'(x) = \frac{1}{2}\|Tx - b\|^2 + \frac{\lambda}{2}\|x\|^2. \tag{14}$$

We then seek a point $x^* \in H_1$ such that

$$J'(x^*) = \inf\{J'(x) : x \in H_1\}. \tag{15}$$

Assuming that J' is Fréchet differentiable at each point of H_1, and given an initial approximation, x_0, the method of steepest descent for minimising $J'(x)$ is now given by

$$x_{n+1} = x_n - \alpha_n \nabla J'(x_n) \tag{16}$$

where $\nabla J'(x_n)$ is the gradient of J' at x_n and the α_n are chosen to minimise $J'(x_{n+1})$ at each step.

The gradient is given by

$$\nabla J'(x) = T^*Tx - T^*b + \lambda x = R' \tag{17}$$

and therefore the steepest descent algorithm becomes

$$x_{n+1} = x_n - \alpha_n R'_n. \tag{18}$$

Choosing α_n to minimise $J'(x_{n+1})$

$$J'(x_{n+1}) = \frac{1}{2}\|Tx_{n+1} - b\|^2 + \frac{\lambda}{2}\|x_{n+1}\|^2$$

but $x_{n+1} = x_n - \alpha_n R'_n$, therefore

$$J'(x_{n+1}) = \frac{1}{2}\|T(x_n - \alpha_n R'_n) - b\|^2 + \frac{\lambda}{2}\|x_n - \alpha_n R'_n\|^2.$$

Substituting $r_n = Tx_n - b$ and expanding

$$\begin{aligned} J'(x_{n+1}) &= \frac{1}{2}\langle r_n - \alpha_n TR'_n, r_n - \alpha_n TR'_n\rangle + \frac{\lambda}{2}\langle x_n - \alpha_n R'_n, x_n - \alpha_n R'_n\rangle \\ &= \frac{1}{2}\langle r_n, r_n\rangle - \alpha_n\langle r_n, TR'_n\rangle + \frac{\alpha_n^2}{2}\langle TR'_n, TR'_n\rangle + \frac{\lambda}{2}\langle x_n, x_n\rangle \\ &\quad -\alpha_n\lambda\langle x_n, R'_n\rangle + \frac{\alpha_n^2}{2}\langle R'_n, R'_n\rangle. \end{aligned}$$

This is minimised for

$$\frac{\partial J'(x_{n+1})}{\partial \alpha_n} = 0.$$

Therefore

$$-\langle r_n, TR'_n\rangle + \alpha_n\langle TR'_n, TR'_n\rangle - \lambda\langle x_n, R'_n\rangle + \alpha_n\langle R'_n, R'_n\rangle = 0$$

from which

$$\alpha_n = \frac{\langle r_n, TR'_n\rangle + \lambda\langle x_n, R'_n\rangle}{\langle TR'_n, TR'_n\rangle + \lambda\langle R'_n, R'_n\rangle}.$$

But

$$\begin{aligned} \langle r_n, TR'_n\rangle = \langle T^* r_n, R'_n\rangle &= \langle R'_n - \lambda x_n, R'_n\rangle \\ &= \langle R'_n, R'_n\rangle - \lambda\langle x_n, R'_n\rangle \end{aligned}$$

and therefore, finally,

$$\alpha_n = \frac{\|R'_n\|^2}{\|TR'_n\|^2 + \lambda\|R'_n\|^2}. \tag{19}$$

As a check we have

$$
\begin{aligned}
\frac{\partial^2 J'(x_{n+1})}{\partial \alpha_n^2} &= \langle TR'_n, TR'_n\rangle + \lambda\langle R'_n, R'_n\rangle \\
&= \|TR'_n\|^2 + \lambda\|R'_n\|^2 \geq 0
\end{aligned}
$$

for $\lambda \geq 0$, therefore the particular choice of α_n, (19), does, in fact, minimise $J'(x_{n+1})$.

The method of steepest descent for minimising the functional $J'(x)$ for $b \in H_2$ is defined, for a given initial approximation, x_0, by the following sequence

$$
x_{n+1} = x_n - \alpha_n R'_n, \quad n = 0, \ldots \tag{20}
$$

where

$$
R'_n = T^*Tx_n - T^*b + \lambda x_n \tag{21}
$$

and

$$
\alpha_n = \frac{\|R'_n\|^2}{\|TR'_n\|^2 + \lambda\|R'_n\|^2}. \tag{22}
$$

Theorem 4.1 *Let H_1 and H_2 be Hilbert spaces and T be a bounded linear operator on H_1 into H_2 such that its range, $\mathcal{R}(T)$, is closed. The sequence of steepest descent defined by (20)-(22) converges to the unique solution, $(T^*T + \lambda I)^{-1}T^*b$, of $(T^*T + \lambda I)x = T^*b$ for any $x_0 \in \mathcal{R}(T^*) = \mathcal{R}(T^*T)$.*

Proof.

$$
\begin{aligned}
J'(x_{n+1}) &= \frac{1}{2}\langle Tx_n - b - \alpha_n TR'_n, Tx_n - b - \alpha_n TR'_n\rangle \\
&\quad + \frac{\lambda}{2}\langle x_n - \alpha_n R'_n, x_n - \alpha_n R'_n\rangle.
\end{aligned}
$$

Expanding

$$
\begin{aligned}
J'(x_{n+1}) &= \frac{1}{2}\|Tx_n - b\|^2 - \alpha_n\langle Tx_n - b, TR'_n\rangle + \frac{\alpha_n^2}{2}\|TR'_n\|^2 + \frac{\lambda}{2}\|x_n\|^2 \\
&\quad - \alpha_n\lambda\langle x_n, R'_n\rangle + \frac{\alpha_n^2}{2}\lambda\|R'_n\|^2.
\end{aligned}
$$

But $J'(x_n) = \frac{1}{2}\|Tx_n - b\|^2 + \frac{\lambda}{2}\|x_n\|^2$, thus

$$
\begin{aligned}
J'(x_{n+1}) &= J'(x_n) - \alpha_n\langle Tx_n - b, TR'_n\rangle + \frac{\alpha_n^2}{2}\|TR'_n\|^2 - \\
&\quad \alpha_n\lambda\langle x_n, R'_n\rangle + \frac{\alpha_n^2}{2}\lambda\|R'_n\|^2
\end{aligned}
$$

which can also be written as

$$
\begin{aligned}
J'(x_{n+1}) &= J'(x_n) - \alpha_n\langle T^*Tx_n - T^*b + \lambda x_n - \lambda x_n, R'_n\rangle + \frac{\alpha_n^2}{2}\|TR'_n\|^2 \\
&\quad - \alpha_n\lambda\langle x_n, R'_n\rangle + \frac{\alpha_n^2}{2}\lambda\|R'_n\|^2.
\end{aligned}
$$

Now, $T^*Tx_n - T^*b + \lambda x_n = R'_n$, and thus

$$J'(x_{n+1}) = J'(x_n) - \alpha_n \langle R'_n - \lambda x_n, R'_n \rangle + \frac{\alpha_n^2}{2}\|TR'_n\|^2 - \alpha_n \lambda \langle x_n, R'_n \rangle + \frac{\alpha_n^2}{2}\lambda \|R'_n\|^2.$$

Therefore

$$\begin{aligned} J'(x_{n+1}) &= J'(x_n) - \alpha_n \|R'_n\|^2 + \lambda \alpha_n \langle x_n, R'_n \rangle + \frac{\alpha_n^2}{2}\|TR'_n\|^2 \\ &\quad -\lambda \alpha_n \langle x_n, R'_n \rangle + \frac{\alpha_n^2}{2}\lambda \|R'_n\|^2 \end{aligned}$$

and finally

$$J'(x_{n+1}) = J'(x_n) - \alpha_n \|R'_n\|^2 + \frac{\alpha_n^2}{2}\|TR'_n\|^2 + \frac{\alpha_n^2}{2}\lambda \|R'_n\|^2.$$

Substituting for α_n, (19),

$$\begin{aligned} J'(x_{n+1}) &= J'(x_n) - \frac{\|R'_n\|^4}{\|TR'_n\|^2 + \lambda \|R'_n\|^2} + \frac{1}{2}\frac{\|R'_n\|^4}{(\|TR'_n\|^2 + \lambda \|R'_n\|^2)^2}\|TR'_n\|^2 \\ &\quad + \frac{\lambda}{2}\frac{\|R'_n\|^4}{(\|TR'_n\|^2 + \lambda \|R'_n\|^2)^2}\|R'_n\|^2 \\ &= J'(x_n) - \frac{\|R'_n\|^4}{\|TR'_n\|^2 + \lambda \|R'_n\|^2} \\ &\quad + \frac{1}{2}\frac{\|R'_n\|^4}{(\|TR'_n\|^2 + \lambda \|R'_n\|^2)^2}(\|TR'_n\|^2 + \lambda \|R'_n\|^2) \end{aligned}$$

and finally

$$J'(x_{n+1}) = J'(x_n) - \frac{1}{2}\frac{\|R'_n\|^4}{\|TR'_n\|^2 + \lambda \|R'_n\|^2}.$$

Therefore $J'(x_{n+1}) \leq J'(x_n)$ for all n, with equality holding when $R'_n = 0$.

Recursively

$$J'(x_{n+1}) = J'(x_0) - \frac{1}{2}\sum_{i=0}^{n}\frac{\|R'_i\|^4}{\|TR'_i\|^2 + \lambda \|R'_i\|^2}.$$

Since $J'(x)$ $(= \frac{1}{2}\|Tx - b\|^2 + \frac{\lambda}{2}\|x\|^2)$ is bounded below by zero

$$\sum_{i=0}^{\infty}\frac{\|R'_i\|^4}{\|TR'_i\|^2 + \lambda \|R'_i\|^2} < \infty. \tag{23}$$

Moreover, by Schwarz's inequality, $\|TR'_i\|^2 \leq \|T\|^2\|R'_i\|^2$ and therefore

$$\begin{aligned} \|TR'_i\|^2 + \lambda \|R'_i\|^2 &\leq \|T\|^2\|R'_i\|^2 + \lambda \|R'_i\|^2 \\ &= \|R'_i\|^2(\|T\|^2 + \lambda). \end{aligned}$$

From which

$$\sum_{i=0}^{\infty} \frac{\|R'_i\|^4}{\|T\|^2\|R'_i\|^2 + \lambda\|R'_i\|^2} = \sum_{i=0}^{\infty} \frac{\|R'_i\|^4}{\|R'_i\|^2(\|T\|^2 + \lambda)}$$
$$= \frac{1}{\|T\|^2 + \lambda} \sum_{i=0}^{\infty} \|R'_i\|^2 \leq \sum_{i=0}^{\infty} \frac{\|R'_i\|^4}{\|TR'_i\|^2 + \lambda\|R'_i\|^2}. \tag{24}$$

Combining (23) and (24)

$$\sum_{i=0}^{\infty} \|R'_i\|^2 < \infty$$

and therefore $R'_n = T^*Tx_n - T^*b + \lambda x_n \to 0$ as $n \to \infty$.

All that remains is to show strong convergence of $\{x_n\}$. By recursion

$$x_{n+1} = x_0 - \sum_{i=0}^{n} \alpha_i R'_i \tag{25}$$

Hence, for $m > n$,

$$x_m - x_n = -\sum_{i=n}^{m-1} \alpha_i R'_i. \tag{26}$$

Now

$$R'_i = T^*Tx_i - T^*b + \lambda x_i = (T^*T + I)x_i - T^*b$$

Therefore, if $x_0 \in \mathcal{R}(T^*)$ we must have $R'_i \in \mathcal{R}(T^*)$ for all i and therefore $x_m - x_n \in \mathcal{R}(T^*)$ for all m, n. Then, by Lemma 1.1

$$\delta^2\|x_m - x_n\|^2 \leq \|T(x_m - x_n)\|^2 = \langle T^*T(x_m - x_n), x_m - x_n\rangle. \tag{27}$$

But

$$\begin{aligned}
&\langle T^*T(x_m - x_n), x_m - x_n\rangle \\
&\quad = \langle T^*T(x_m - x_n) - T^*b + T^*b - \lambda x_m + \lambda x_m - \lambda x_n + \lambda x_n, x_m - x_n\rangle \\
&\quad = \langle T^*Tx_m - T^*b + \lambda x_m, x_m - x_n\rangle - \langle T^*Tx_n - T^*b + \lambda x_n, x_m - x_n\rangle - \\
&\qquad \lambda\langle x_m - x_n, x_m - x_n\rangle
\end{aligned}$$

and therefore

$$\begin{aligned}
\delta^2\|x_m - x_n\|^2 \leq\ & \langle T^*Tx_m - T^*b + \lambda x_m, x_m - x_n\rangle - \\
& \langle T^*Tx_n - T^*b + \lambda x_n, x_m - x_n\rangle - \lambda\|x_m - x_n\|^2.
\end{aligned}$$

Since $\lambda \geq 0$ we have $\delta^2 + \lambda > 0$ and therefore

$$\begin{aligned}
&(\delta^2 + \lambda)\|x_m - x_n\|^2 \\
&\quad \leq \langle T^*Tx_m - T^*b + \lambda x_m, x_m - x_n\rangle - \langle T^*Tx_n - T^*b + \lambda x_n, x_m - x_n\rangle \\
&\quad \leq |\langle T^*Tx_m - T^*b + \lambda x_m, x_m - x_n\rangle| + |\langle T^*Tx_n - T^*b + \lambda x_n, x_m - x_n\rangle| \\
&\quad \leq \|T^*Tx_m - T^*b + \lambda x_m\|\|x_m - x_n\| + \|T^*Tx_n - T^*b + \lambda x_n\|\|x_m - x_n\|
\end{aligned}$$

But, $\|x_m - x_n\| \leq 1/\delta \|T(x_m - x_n)\|$, and therefore

$$(\delta^2 + \lambda)\|x_m - x_n\|^2 \leq$$
$$(1/\delta)(\|T^*Tx_m - T^*b + \lambda x_m\| + \|T^*Tx_n - T^*b + \lambda x_n\|)\|T(x_m - x_n)\|.$$

$\{T(x_m - x_n)\}$ is bounded, say $\|T(x_m - x_n)\| \leq M$ and hence

$$\begin{aligned} &(\delta^2 + \lambda)\|x_m - x_n\|^2 \\ \leq\ & \frac{M}{\delta}(\|T^*Tx_m - T^*b + \lambda x_m\| + \|T^*Tx_n - T^*b + \lambda x_n\|) \\ =\ & \frac{M}{\delta}(R'_m + R'_n) \rightarrow 0 \text{ as } m, n \rightarrow \infty. \end{aligned}$$

Thus $\{x_n\}$ is a Cauchy sequence and therefore converges to an element, $u \in H_1$, and

$$\lim_{n \to \infty} J'(x_n) = J'(u) = \inf\{J'(x) : x \in H_1\} \tag{28}$$

Since

$$T^*Tx_n - T^*b + \lambda x_n = R'_n \rightarrow 0 \tag{29}$$

then

$$(T^*T + \lambda I)u = T^*b \tag{30}$$

or

$$u = (T^*T + \lambda I)^{-1}T^*b \tag{31}$$

i.e. u is the unique solution (Corollary 2.1).

Now, we have already required that $x_0 \in \mathcal{R}(T^*)$ (to ensure $R'_i \in \mathcal{R}(T^*)$) and therefore

$$x_{n+1} = x_0 - \sum_{i=0}^{n} \alpha_i R'_i \in \mathcal{R}(T^*) \tag{32}$$

since $R'_i \in \mathcal{R}(T^*)$ for all i. Since $\mathcal{R}(T^*)$ is closed then $u \in \mathcal{R}(T^*)$. Now $(T^*T + \lambda I)^{-1}T^*b$ is the unique solution (in $\mathcal{R}(T^*)$) and therefore we must have that $\{x_n\}$ converges to $u = (T^*T + \lambda I)^{-1}T^*b$. $\square$

5 Application to Tikhonov Regularisation

The problem of solving (6) is said to be well posed if a unique solution exists which depends continuously on b. Following (de Mol 1992) we adopt a less restrictive definition since we know that a unique generalised solution exists. The following statements are equivalent:

(i) the problem of solving $Tx = b$ is well posed;

(ii) $x^\dagger$ exists for any $b \in H_2$;

(iii) $T^\dagger$ is continuous (equivalently bounded);

(iv) $\mathcal{R}(T)$ is closed; and

(v) $\lambda = 0$ is not an accumulation point of the spectrum of T^*T.

Otherwise, the problem is said to be ill-posed, i.e. the generalised inverse, $T^\dagger$, is unbounded. In order to obtain estimates of $x^\dagger$ which are stable to variations in b we must seek regularised solutions of the ill-posed problem.

We have restricted our attention to the case where $\mathcal{R}(T)$ is closed (and hence $T^\dagger$ bounded), for which the problem of solving (6) is always well posed. However, this does not ensure that the generalised solution will be numerically stable. The relative error in the generalised solution corresponding to variations in b is bounded as follows:

$$\frac{||\delta x^\dagger||}{||x^\dagger||} \leq C(T)\frac{||\delta b||}{||b||} \tag{33}$$

where $C(T)$ is the condition number of the operator T, given by

$$C(T) = ||T||||T^\dagger|| = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}. \tag{34}$$

λ_{min} and λ_{max} are, respectively, the lower and upper limits of the positive part of the spectrum of TT^*. For a large condition number, the amplification of δb may cause the generalised solution, $x^\dagger$, to be unstable. The problem is then said to be ill-conditioned and, even though well-posed, we still need to apply regularisation to achieve a stable solution. This is the reason for discussing regularisation of the linear operator equation, (6), even though $\mathcal{R}(T)$ is closed.

We now describe the particular case of the Tikhonov regularisation method (Groetsch 1984). Consider the problem of minimising the functional (c.f. (14))

$$\Phi_\lambda(x) = \frac{1}{2}||Tx - b||^2 + \frac{\lambda}{2}||x||^2. \tag{35}$$

The quadratic functional $\Phi_\lambda(x)$ has a unique minimum, denoted by x_λ, which is a solution of

$$(T^*T + \lambda I)x_\lambda = T^*b \tag{36}$$

where I denotes the appropriate identity operator. We have already seen that the operator $(T^*T + \lambda I)$ is invertible for $\lambda > 0$ and hence

$$x_\lambda = L_\lambda b, \quad \text{where } L_\lambda = (T^*T + \lambda I)^{-1}T^* \tag{37}$$

is known as the Tikhonov regularisation operator. We can also write the regulariser as follows

$$L_\lambda = T^*(TT^* + \lambda I)^{-1} \tag{38}$$

from which it is clear that $x_\lambda \in \mathcal{R}(T^*) \subset N(T)^\perp$ and hence (37) defines a regularisation method (for a proof see (Groetsch 1984; Kirsch 1996)).

The following is therefore a natural corollary of the method of steepest descent, Theorem 4.1, which provides an iterative solution to the Tikhonov regularisation method.

Corollary 5.1 *Let H_1 and H_2 be Hilbert spaces and T be a bounded linear operator on H_1 into H_2 such that its range, $\mathcal{R}(T)$, is closed. Assume, further, that the problem of solving the linear operator equation, $Tx = b, b \in H_2$ for $x \in H_1$ is ill-conditioned. The sequence of steepest descent defined by (20)-(22) converges to the unique Tikhonov regularised solution, $(T^*T + \lambda I)^{-1}T^*b$ for any $x_0 \in \mathcal{R}(T^*) = \mathcal{R}(T^*T)$.*

Acknowledgements

The authors would like to thank the UK EPSRC for their financial support under Grant No. GR/R15726/01.

References

Ben-Israel, A. and A. Charnes (1963). Contribution to the theory of generalized inverses. *Journal of the Society for Industrial and Applied Mathematics 11*(3), 667–699.

de Mol, C. (1992). A critical survey of regularized inversion methods. In M. Bertero and E. Pike (Eds.), *Inverse Problems in Scattering and Imaging, Proceedings of the NATO Advanced Research Workshop held at Cape Cod, USA, 14-19 April 1991*, pp. 345–370. IOP Publishing.

Desoer, C. and B. Whalen (1963). A note on pseudoinverses. *Journal of the Society for Industrial and Applied Mathematics 11*(2), 442–447.

Groetsch, C. (1977). *Generalized Inverses of Linear Operators.* Monographs and Textbooks in Pure and Applied Mathematics. Marcel Dekker.

Groetsch, C. (1984). *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Volume 105 of *Research Notes in Mathematics Series.* Pitman.

Kammerer, W. and M. Nashed (1972). Iterative methods for best approximate solutions of linear integral equations of the first and second kinds. *Journal of Mathematical Analysis and Applications 40*, 547–573.

Kantorovich, L. and G. Akilov (1964). *Functional Analysis in Normed Spaces*, Volume 46 of *International Series of Monographs in Pure and Applied Mathematics.* Pergamon Press.

Kirsch, A. (1996). *An Introduction to the Mathematical Theory of Inverse Problems*, Volume 120 of *Applied Mathematical Sciences.* Springer.

McCormick, S. and G. Rodrigue (1975). A uniform approach to gradient methods for linear operator equations. *Journal of Mathematical Analysis and Applications 49*, 275–285.

Nashed, M. (1970). Steepest descent for singular linear operator equations. *SIAM Journal on Numerical Analysis 7*(3), 358–362.

Nashed, M. (1971). Generalized inverses, normal solvability, and iteration for singular operator equations. In L. B. Rall (Ed.), *Nonlinear Functional Analysis and Applications, Proceedings of an Advanced Seminar Conducted by the Mathematics Research Center, The University of Wisconsin, Madison, October 12-14, 1970*, pp. 311–359. Academic Press.

Nashed, M. Z. (Ed.) (1976). *Generalized Inverses and Applications.* Proceedings of an Advanced Seminar Conducted by the Mathematics Research Center, The University of Wisconsin, Madison, October 12-14, 1970: Academic Press.

Petryshyn, W. (1967). On generalized inverses and on the uniform convergence of $(I - \beta K)^n$ with applications to iterative methods. *Journal of Mathematical Analysis and Applications 18*, 417–439.