Proc. Estonian Acad. Sci. Phys. Math., 2004, **53**, 2, 77–83
https://doi.org/10.3176/phys.math.2004.2.02

On the choice of the regularization parameter for solving self-adjoint ill-posed problems with the approximately given noise level of data

Uno Hämarik and Toomas Raus

Institute of Applied Mathematics, University of Tartu, Liivi 2, 50409 Tartu, Estonia; uno.hamarik@ut.ee, toomas.raus@mtk.ut.ee

Received 20 November 2003

Abstract. We consider ill-posed problems $Au = f$, where the operator $A \in \mathcal{L}(H, H)$, $A = A^* \geq 0$, has a nonclosed range in the Hilbert space H. We assume that instead of f noisy data $\tilde{f}$ are given, with the approximately known noise level δ. The problem $Au = f$ is regularized by the (iterated) Lavrentiev method, by iterative methods or by the method of the Cauchy problem. For the choice of the regularization parameter we propose a new *a posteriori* rule with the property that the regularized solution converges to the exact one in the process $\delta \to 0$ provided that the ratio $\|\tilde{f} - f\|/\delta$ is bounded for $\delta \to 0$. The error estimates are given.

Key words: ill-posed problems, regularization methods, Lavrentiev method, iterative method, method of the Cauchy problem, noise level, parameter choice.

1. INTRODUCTION

We consider an operator equation

$$Au = f\,, \quad f \in \mathcal{R}(A), \tag{1}$$

with the linear continuous self-adjoint and non-negative operator $A \in \mathcal{L}(H, H)$, $A = A^* \geq 0$, and u and f being elements of the real Hilbert space H. We do not suppose that the range $\mathcal{R}(A)$ is closed and so in general our problem is ill-posed. The kernel $N(A)$ may be nontrivial. We suppose that instead of the exact right-hand side f we have an approximation $\tilde{f} \in H$ with the supposable noise level δ, but it is unknown whether the inequality $\|\tilde{f} - f\| \leq \delta$ holds or not.

For a stable numerical approximation of the solution of (1) some regularization technique is needed. Here we consider the following regularization methods.

1. The Lavrentiev method $u_\alpha = (\alpha I + A)^{-1}(\alpha u_0 + \tilde{f})$.

2. The iterated Lavrentiev method. Let $u_0 = u_{0,\alpha} \in H$ be the initial approximation, find $u_{k,\alpha} = (\alpha I + A)^{-1}(\alpha u_{k-1,\alpha} + \tilde{f})$ for $k = 1, \ldots, m$, and take $u_\alpha = u_{m,\alpha}$.

3. The explicit iteration scheme (Landweber's method)

$$u_n = u_{n-1} - \mu(Au_{n-1} - f)\,, \quad n = 1, 2, \ldots\ (0 < \mu < 1/\|A\|)\,.$$

4. The implicit iteration scheme $\alpha u_n + Au_n = \alpha u_{n-1} + \tilde{f}$, $n = 1, 2, \ldots$ $(\alpha = \mathrm{const} > 0)$.

5. The method of the Cauchy problem. We take for u_r the solution of the Cauchy problem

$$u'(r) + Au(r) = \tilde{f}\,, \quad u(0) = u_0\,.$$

All these methods can be presented in the general form (see [1,2])

$$u_r = (I - Ag_r(A))u_0 + g_r(A)\tilde{f}\,, \tag{2}$$

where u_r is the approximate solution, u_0 – initial approximation, r – regularization parameter (in methods 1 and 2, $r = 1/\alpha$; in methods 3 and 4, $r = n$), I – identity operator, and the function $g_r(\lambda)$ satisfies the conditions (3) and (4):

$$\sup_{0 \le \lambda \le a} \big|g_r(\lambda)\big| \le \gamma r\,, \quad r \ge 0\,, \tag{3}$$

$$\sup_{0 \le \lambda \le a} \lambda^p \big|1 - \lambda g_r(\lambda)\big| \le \gamma_p r^{-p}\,, \quad r \ge 0\,, 0 \le p \le p_0\,. \tag{4}$$

Here p_0, γ, and γ_p are positive constants, $\gamma_0 \le 1$, $a \ge \|A\|$ and the greatest value of p_0, for which the inequality (4) holds, is called the qualification of the method. For methods 1–5 the corresponding generating functions are

$$g_r(\lambda) = \frac{r}{1 + r\lambda}, \quad g_r(\lambda) = \frac{1}{\lambda}\Big(1 - \Big(\frac{1}{1 + r\lambda}\Big)^m\Big),$$

$$g_r(\lambda) = \frac{1}{\lambda}\big(1 - (1 - \mu\lambda)^r\big), \quad g_r(\lambda) = \frac{1}{\lambda}\Big(1 - \Big(\frac{\alpha}{\alpha + \lambda}\Big)^r\Big), \quad g_r(\lambda) = \frac{1}{\lambda}\big(1 - e^{-r\lambda}\big).$$

The qualifications of methods 1 and 2 are $p_0 = 1$ and $p_0 = m$, respectively; in methods 3–5, $p_0 = \infty$.

2. KNOWN RULES FOR THE CHOICE OF THE REGULARIZATION PARAMETER

By using the regularization method (2) with noisy data an important problem is the choice of an appropriate regularization parameter r. For the choice of the parameter r two types of rules can be found in literature. Rules of the first type use information about the noise level of data $\tilde{f}$, rules of the second type do not use such information. Below, we consider two rules of **the first group**.

1. Morozov's discrepancy principle [3]. In this rule the regularization parameter $r = r_{\mathrm{D}}$ is chosen as the solution of the equation

$$\big\|Au_r - \tilde{f}\big\| \approx b\delta \quad \text{with} \quad b = \text{const} > 1\,.$$

2. Modification of the discrepancy principle [4]. In this rule the regularization parameter $r = r_{\mathrm{MD}}$ is chosen as the solution of the equation

$$\big\|B_r\big(Au_r - \tilde{f}\big)\big\| \approx b\delta \quad \text{with} \quad b = \text{const} > 1\,,$$

   where the operator $B_r = I$ for $p_0 = \infty$, $B_r = (I - Ag_r(A))^{1/p_0}$ for $p_0 \neq \infty$. Note that for the Lavrentiev method and its iterative variant $\|B_r(Au_{\alpha,m} - \tilde{f})\| = \|Au_{\alpha,m+1} - \tilde{f}\|$.

Some useful properties of the modified discrepancy principle are (see [4])

1. convergence: $\|u_{r_{\mathrm{MD}}} - u_*\| \to 0$ for $\delta \to 0$; here u_* is nearest to u_0 solution of the problem $Au = f$;

2. order-optimality: if $u_0 - u_* = A^p v$, $v \in H$, $\|v\| \leq \varrho$, $p > 0$, then $\|u_{r_{\mathrm{MD}}} - u_*\| \leq C_p \varrho^{1/(p+1)} \delta^{p/(p+1)}$, $0 < p \leq p_0$;

3. quasioptimality: there exists a constant c such that

$$\|u_{r_{\mathrm{MD}}} - u_*\| \leq c \inf_{r \geq 0} \{\|(I - Ag_r(A))(u_0 - u_*)\| + \gamma r\delta\}\,.$$

However, the discrepancy principle and its modification have an essential disadvantage. Namely, these choices are unstable in the sense that if the actual error of the right-hand side is larger than $b\delta$, the error of the approximate solution may be arbitrarily large independently of the value of the ratio of the actual to the supposed noise level.

The parameter choice rules of **the second group** are sometimes called heuristic or delta-free choices. The first rule of this kind was the quasioptimality criterion [5], in which such a parameter is chosen for which the function $k(r) = r\|B_r(Au_r - \tilde{f})\|$ has the minimum. Other popular delta-free rules are Wahba's cross-validation rule [6] and Hansen's L-curve rule [7]. Some heuristic rules are also proposed in [8].

Although these rules often work well, it was shown by Bakushinskii [9] that the convergence of the approximate solution for heuristic parameter choice rules cannot be proved.

3. PARAMETER CHOICE RULE IN THE CASE OF THE APPROXIMATELY GIVEN NOISE LEVEL OF DATA

In practice it may be complicated to find the exact noise level, but we often have a rough guess at it. In the following we assume that the supposed error level $\delta > 0$ is given, but we do not know exactly if $\|\tilde{f} - f\| \leq \delta$ holds or not. Our aim is to present a rule for the stable parameter choice which guarantees the convergence of the approximate solution to the exact solution if only the ratio $\|\tilde{f} - f\|/\delta$ is bounded in the process $\delta \to 0$, and to give some error estimates of the approximate solution.

In the following the function

$$\varphi(r) = \sqrt{r}\big\|A^{1/2}B_r^{3/2}\big(Au_r - \tilde{f}\big)\big\| = \sqrt{r}\big\langle B_r\big(Au_r - \tilde{f}\big), AB_r^2\big(Au_r - \tilde{f}\big)\big\rangle^{1/2}$$

plays an important role.

Note that for the Lavrentiev method and its iterative variant $B_r = (I + rA)^{-1}$ and $\varphi(r) = \varphi(\alpha^{-1}) = \frac{1}{\sqrt{\alpha}}\langle Au_{m+1,\alpha} - \tilde{f}, A(Au_{m+2,\alpha} - \tilde{f})\rangle^{1/2}$; for iterative methods $\varphi(r) = \varphi(n) = \sqrt{n}\langle Au_n - \tilde{f}, Au_n - \tilde{f}\rangle^{1/2}$.

Rule P. Let $0 \leq s \leq 1$ and b_1, b_2 be the constants such that $b_2 \geq b_1 > C_m$, where $C_m = 1/2$, $C_m = 1/\sqrt{2m+3}$, $C_m = 1/\sqrt{2\mu e}$, $C_m = \sqrt{\alpha/2}$, and $C_m = 1/\sqrt{2e}$ for methods 1–5, respectively. If $\varphi(1) \leq b_2\delta$, choose $r(\delta) = 1$. In the opposite case we find at first $r_2(\delta) > 1$ such that

$$\varphi(r_2(\delta)) \leq b_2\delta\,, \tag{5}$$

$$\varphi(r) \geq b_1\delta \qquad \forall\, r \in [1, r_2(\delta)]\,. \tag{6}$$

As the regularization parameter $r(\delta)$ we choose the parameter r, for which the function $t(r) = r^s\|B_r(Au_r - \tilde{f})\|$ has the global minimum on the interval $[1, r_2(\delta)]$.

In [10] the parameter choice rule was considered, in which $r_2(\delta)$ was taken as the regularization parameter. We can regard Rule P as the generalization of this rule, since in case $s = 0$ these rules coincide due to the fact that the function $\|B_r(Au_r - \tilde{f})\|$ is monotonically decreasing with respect to r. On the other hand, in case $s = 1$ Rule P is similar to the parameter choice by the quasioptimality criterion, where the minimum point of the function $r\|B_r(Au_r - \tilde{f})\|$ is chosen as the regularization parameter. The difference between Rule P and the quasioptimality criterion is the interval, on which the function $r\|B_r(Au_r - \tilde{f}\|$ is minimized: the intervals are $[1, r_2(\delta)]$ and $[1, \infty)$, respectively.

In [10] it was proven for methods 1–5 that for each $\tilde{f} \in H$ we have $\lim_{r\to\infty} \varphi(r) = 0$. Due to the continuity of the function $\varphi(r)$, it follows from this property that the choice of the finite parameters $r_2(\delta)$ and $r(\delta) \le r_2(\delta)$ according to Rule P is possible. Note that the function $\varphi(r)$ is non-monotone and therefore in Rule P we must use the conditions (5)–(6) instead of the inequalities $b_1\delta \le \varphi(r) \le b_2\delta$.

Theorem 1. *Let* $A \in L(H, H)$, $A = A^* \ge 0$, $f \in \mathcal{R}(A)$. *Let the parameter* $r(\delta)$ *be chosen according to Rule P. If* $\frac{\|\tilde{f}-f\|}{\delta} \le \text{const}$ *in the process* $\delta \to 0$, *then for methods* 1–5, $\left\|u_{r(\delta)} - u_*\right\| \to 0$.

Proof. Denote $G_r := I - Ag_r(A)$. Then we have $u_r - u_* = G_r(u_0 - u_*) + g_r(A)(\tilde{f} - f)$, and (3) yields

$$\left\|u_{r(\delta)} - u_*\right\| \le \left\|G_{r(\delta)}(u_0 - u_*)\right\| + \gamma C r(\delta)\delta\,. \tag{7}$$

In [10] it is proven that $r_2(\delta)\delta \to 0$ for $\delta \to 0$. This convergence and the inequality $r(\delta) \le r_2(\delta)$ yield the convergence of the second term of (7).

From the equality

$$B_r(Au_r - \tilde{f}) = AB_rG_r(u_0 - u_*) - B_rG_r(\tilde{f} - f) \tag{8}$$

and from the inequality $\|B_rG_r(\tilde{f} - f)\| \le \|\tilde{f} - f\| \le C\delta$ it follows that

$$r_2^s(\delta)\left\|B_{r_2(\delta)}(Au_{r_2(\delta)} - \tilde{f})\right\| \le r_2^s(\delta)\left\|AB_{r_2(\delta)}G_{r_2(\delta)}(u_0 - u_*)\right\| + r_2^s(\delta)C\delta\,. \tag{9}$$

To show the convergence

$$r_2^s(\delta)\left\|AB_{r_2(\delta)}G_{r_2(\delta)}(u_0 - u_*)\right\| \to 0 \quad \text{for} \quad \delta \to 0\,, \tag{10}$$

we consider separately the cases a) $r_2(\delta) \to \infty$ and b) $r_2(\delta) \le \overline{r} = \text{const}$. If $r_2(\delta) \to \infty$ in the process $\delta \to 0$, then using the Banach–Steinhaus theorem, we can prove similarly to [2] (p. 43) that $r^p\|AB_rG_r(u_0 - u_*)\| \to 0$ if $r \to \infty$ $(0 \le p \le 1)$. Now we consider the case $r_2(\delta) \le \overline{r} = \text{const}$. Using (8), (4), we get

$$\begin{aligned} &r_2(\delta)^{1/2}\left\|A^{3/2}B^{3/2}_{r_2(\delta)}G_{r_2(\delta)}(u_0 - u_*)\right\| \\ &\quad\le r_2(\delta)^{1/2}\left\|A^{1/2}B^{3/2}_{r_2(\delta)}(Au_{r_2(\delta)} - \tilde{f})\right\| + r_2(\delta)^{1/2}\left\|A^{1/2}B^{3/2}_{r_2(\delta)}G_{r_2(\delta)}(\tilde{f} - f)\right\| \\ &\quad\le b_2\delta + \gamma_{1/2}\left\|\tilde{f} - f\right\| \le (b_2 + C\gamma_{1/2})\delta, \end{aligned}$$

from which it follows that

$$\left\|A^{3/2}B^{3/2}_{r_2(\delta)}G_{r_2(\delta)}(u_0 - u_*)\right\| \to 0 \quad \text{for} \quad \delta \to 0\,.$$

In [2] (p. 66) the implication

$$AG_{r_n}(u_0 - u_*) \to 0 \ (n \to \infty) \quad \Longrightarrow \quad G_{r_n}(u_0 - u_*) \to 0 \ (n \to \infty) \tag{11}$$

is proven. Similarly we can show that if $A^{3/2}B_{r_n}^{3/2}G_{r_n}(u_0 - u_*) \to 0$ ($n \to \infty$), then $AB_{r_n}G_{r_n}(u_0 - u_*) \to 0$ ($n \to \infty$), which proves the convergence (10) in case $r_2(\delta) \le \mathrm{const}$. From (9), (10) and from the convergence $r_2(\delta)\delta \to 0$ ($\delta \to 0$) it follows that $r_2^s(\delta)\big\|B_{r_2(\delta)}(Au_{r_2(\delta)} - \tilde{f})\big\| \to 0$ for $\delta \to 0$. This convergence yields also the convergence $r^s(\delta)\big\|B_{r(\delta)}(Au_{r(\delta)} - \tilde{f})\big\| \to 0$ for $\delta \to 0$, while $r(\delta)$ is the global minimum point of the function $t(r) = r^s\|B_r(Au_r - \tilde{f})\|$ in $[1, r_2(\delta)]$. Using (8), we get

$$r^s(\delta)\big\|AB_{r(\delta)}G_{r(\delta)}(u_0 - u_*)\big\| \le r^s(\delta)\big\|B_{r(\delta)}(Au_{r(\delta)} - \tilde{f})\big\| + r^s(\delta)C\delta \to 0$$

for $\delta \to 0$, and the implication of type (11) yields $\|G_{r(\delta)}(u_0 - u_*)\| \to 0$ for $\delta \to 0$, which with (7) proves the theorem.

Theorem 2. *Let $A \in L(H, H)$, $A = A^* \ge 0$, $f \in \mathcal{R}(A)$, $\|\tilde{f} - f\| \le \delta$. Let the parameter $r(\delta)$ be chosen according to Rule P with $s \in (0, 1)$ and let the function $t(r) = r^s\|B_r(Au_r - \tilde{f})\|$ be monotonically increasing on the interval $[r(\delta), r_2(\delta)]$. Then for methods* 1–5 *the error estimation*

$$\big\|u_{r(\delta)} - u_*\big\| \le c(b_1, b_*)\frac{1}{1-s}\inf_{r \ge 0}\left\{\big\|(I - Ag_r(A))(u_0 - u_*)\big\| + \gamma r\delta\right\}$$

holds, where $b_* = \max_{r(\delta) \le r \le R(\delta)} \varphi(r)/\delta \ge b_2$ and $R(\delta)$ is the greatest parameter for which $\varphi(r) = b_2\delta$.

Theorem 3. *Let $A \in L(H, H)$, $A = A^* \ge 0$, $f \in \mathcal{R}(A)$. Let the parameter $r(\delta)$ be chosen according to Rule P with $s \in (0, 1)$. Then in case $\|\tilde{f} - f\| > \delta$ for methods* 1–5 *the following error estimations hold*:

a) if $\delta \le \|\tilde{f} - f\| \le \delta_0$, where $\delta_0 := \|B_{r(\delta)}(Au_{r(\delta)} - \tilde{f})\|$, then

$$\big\|u_{r(\delta)} - u_*\big\| \le c_1(b_1, b_*)\frac{1}{1-s}\inf_{r \ge 0}\left\{\big\|(I - Ag_r(A))(u_0 - u_*)\big\| + \gamma r\big\|\tilde{f} - f\big\|\right\};$$

b) if $\|\tilde{f} - f\| > \delta_0$, then

$$\big\|u_{r(\delta)} - u_*\big\| \le c_2(b_1, b_2)\left(\frac{\|\tilde{f} - f\|}{\delta_0}\right)^{1/s}\inf_{r \ge 0}\left\{\big\|(I - Ag_r(A))(u_0 - u_*)\big\| + \gamma r\big\|\tilde{f} - f\big\|\right\}.$$

Note that all coefficients $c(b_1, b_*)$, $c_1(b_1, b_*)$, $c_2(b_1, b_2)$ in methods 1–3, 5 are less than 2.5 in case $b_1 = b_2 = 1.5C_m$, $b_* = b_2$.

The proofs of Theorems 2, 3 will be presented in a forthcoming paper.

ACKNOWLEDGEMENT

This work was supported by the Estonian Science Foundation (grant No. 5785).

REFERENCES

1. Vainikko, G. The discrepancy principle for a class of regularization methods. *U.S.S.R. Comput. Math. Math. Phys.*, 1982, **22**, 1–19.
2. Vainikko, G. and Veretennikov, A. *Iteration Procedures in Ill-Posed Problems*. Nauka, Moscow, 1986 (in Russian).
3. Morozov, V. On the solution of functional equations by the method of regularization. *Soviet Math. Dokl.*, 1966, **7**, 414–417.
4. Raus, T. Residue principle for ill-posed problems. *Acta Comment. Univ. Tartuensis*, 1984, 672, 16–26 (in Russian).
5. Tikhonov, A. and Arsenin, V. *Solution of Ill-Posed Problems*. Wiley, New York, 1977.
6. Wahba, G. Practical approximate solutions to linear operator equations when the data are noisy. *SIAM J. Numer. Anal.*, 1977, **14**, 651–657.
7. Hansen, P. C. Analysis of discrete ill-posed problems by means of the *L*-curve. *SIAM Rev.*, 1992, **34**, 561–580.
8. Hanke, M. and Raus, T. A general heuristic for choosing the regularization parameter in ill-posed problems. *SIAM J. Sci. Comput.*, 1996, **17**, 954–972.
9. Bakushinskii, A. Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion. *U.S.S.R. Comput. Math. Math. Physics.*, 1984, **24**, 181–182.
10. Raus, T. An a posteriori choice of the regularization parameter in case of approximately given error bound of data. *Acta Comment. Univ. Tartuensis*, 1990, 913, 73–87.

Regulariseerimisparameetri valikust enesekaassete mittekorrektsete ülesannete lahendamisel lähteandmete ligikaudu teadaoleva veataseme korral

Uno Hämarik ja Toomas Raus

On vaadeldud enesekaasse mittenegatiivse operaatoriga mittekorrektse ülesande lahendamist (itereeritud) Lavrentjevi meetodiga, iteratsioonimeetoditega ja Cauchy ülesande meetodiga juhul, kui ülesande lähteandmete veatase on teada ligikaudu. Pakutakse välja regulariseerimisparameetri valiku uus reegel. Tõestatakse lähislahendi koonduvus ja antakse veahinnangud.