Mathematical Modelling and Analysis 2005. Pages 285–291
Proceedings of the 10th International Conference MMA2005&CMAM2, Trakai
 ISBN 9986-05-924-0

COMPARISON OF STOPPING RULES IN CONJUGATE GRADIENT TYPE METHODS FOR SOLVING ILL-POSED PROBLEMS[1]

U. HÄMARIK and R. PALM

University of Tartu

Liivi 2, 50409 Tartu, Estonia

E-mail: `uno.hamarik@ut.ee; reimo.palm@ut.ee`

Abstract. We consider solving linear ill-posed problems by conjugate gradient type methods. Typically only noisy data are available. It is important to stop iterations at the right moment, choosing the stopping index properly according to the noise level. For finding stopping index we formulate the monotone error rule and another rule which works well also for approximately given noise level. Numerical comparison with known rules shows that the new rules are competitive.

Key words: Ill-posed problems, conjugate gradient type methods, noise level, stopping rule, discrepancy principle, monotone error rule

1. Introduction

We consider an operator equation

$$Au = f_* , \qquad f_* \in \mathcal{R}(A), \tag{1.1}$$

where A is a linear continuous operator between Hilbert spaces H and F. In general, the problem (1.1) is ill-posed (see [3, 15]): the range $\mathcal{R}(A)$ may be non-closed, the kernel $\mathcal{N}(A)$ may be non-trivial. In practice often instead of the exact data f_* only an approximation f is given (containing, for example, measurement errors). If an ill-posed problem is solved by some iterative method, then for preventing unbounded magnification of the data error, the iterations should be stopped after a certain number n of steps. If the exact noise level δ with $\|f_* - f\| \leq \delta$ is given, the proper choice of $n = n(\delta)$ guarantees the convergence $u_{n(\delta)} \to u_*$ as $\delta \to 0$, where u_n is an approximation found after n iterations and u_* is the minimal-norm solution of equation (1.1). For many iterative methods this convergence is guaranteed by choice of n by

[1] This work was supported by the Estonian Science Foundation (Grant No. 5785)

the discrepancy principle or by its modifications (see [1,3,4,7,8,10–13,15]) or by the monotone error rule [1,6].

If there is no information about the noise level δ, then no rule can guarantee the convergence $u_{n(\delta)} \to u_*$ $(\delta \to 0)$ (see [2]). Nevertheless, iterations may be stopped e.g. by rules from [9]. In some applications the noise level δ is given approximately: it holds

$$\|f_* - f\|/\delta \le C \quad \text{for } \delta \to 0,$$

where C is an unknown constant. In this case convergence $u_{n(\delta)} \to u_*$ $(\delta \to 0)$ for iterative methods of Landweber and Lardy is guaranteed by stopping iterations by rule from [5].

In this article we formulate the monotone error rule and the analogue of rule [5] for conjugate gradient type methods. We compare these rules and various known stopping rules numerically.

2. Conjugate Gradient Type Methods

The problem (1.1) can be solved by various different iterative methods. The regularizing properties of simple iterative methods (Landweber and Lardy method) are analysed in [15]. The conjugate gradient type methods [1,3,4,7–13] are much faster and much more powerful. In this paper we consider two iterative methods based on conjugate gradient method for various possibilities to symmetrize the problem (1.1). If the conjugate gradient method is applied to the normal equation

$$A^*Au = A^*f,$$

or to the equation

$$AA^*w = f, \quad u = A^*w,$$

we get the methods called CGLS or CGME respectively. In iterative method CGLS the kth iterate u_k minimizes the residual $f - Au$ among all u from the Krylov subspace $\mathrm{span}\{A^*f, A^*AA^*f, \ldots, (A^*A)^{k-1}A^*f\}$ (as in the projection method of least squares). The kth iterate u_k in method CGME minimizes the error $\|u_* - u\|$ with u in the same Krylov subspace (as in the projection method of minimal error, see [8,14]). In both algorithms we fix the starting values $u_0 = 0$, $r_0 = f$, $v_{-1} = 0$. In CGLS we also take $p_{-1} = \infty$ and we compute for every $n = 0, 1, 2, \ldots$

$$\begin{aligned}
&p_n = A^*r_n\,, \quad \sigma_n = \|p_n\|^2/\|p_{n-1}\|^2\,, \quad v_n = r_n + \sigma_n v_{n-1}\,,\\
&q_n = A^*v_n\,, \quad s_n = Aq_n\,, \quad \beta_n = \|p_n\|^2/\|s_n\|^2\,,\\
&u_{n+1} = u_n + \beta_n q_n\,, \quad r_{n+1} = r_n - \beta_n s_n\,.
\end{aligned}$$

In CGME method we take $r_{-1} = \infty$ and compute for every $n = 0, 1, 2, \ldots$

$$\begin{aligned}
&\sigma_n = \|r_n\|^2/\|r_{n-1}\|^2\,, \quad v_n = r_n + \sigma_n v_{n-1}\,, \quad q_n = A^*v_n\,,\\
&\beta_n = \|r_n\|^2/\|q_n\|^2\,, \quad u_{n+1} = u_n + \beta_n q_n\,, \quad r_{n+1} = r_n - \beta_n Aq_n\,.
\end{aligned}$$

3. Stopping Rules

1) First, we consider the case, when the exact noise level δ is known. Then the most prominent stopping rule is the discrepancy principle: we stop at the first index $n = n_D$ for which the value of the function $d_D(n)$ is smaller than $C\delta$, where $C > 1$ is a constant. Function $d_D(n) = \|r_n\|$ is used in the CGLS method and $d_D(n) = \big[\sum\limits_{i=0}^{n} \|r_i\|^{-2}\big]^{-1/2}$ in the CGME method, where $r_n = f - Au_n$. This rule was formulated and studied for the CGLS method in [1, 3, 4, 7, 10–13] and for the method CGME in [7]. As observed in [7], in the method CGLS both functions $d_D(n)$ give the same stopping index n_D. In these works the convergence $\|u_{n_D} - u_*\| \to 0$ $(\delta \to 0)$ was proved and for case $u_* \in \mathcal{R}((A^*A)^{p/2})$ the order optimal error estimate

$$\|u_{n_D} - u_*\| \le c\delta^{p/(p+1)} \text{ for all } p < \infty$$

was stated.

The second rule which we consider, is the monotone error rule. In [1] the stopping index $n = n_{MA}$ in method CGLS is found as the first index for which

$$d_{MA}(n) \equiv \frac{\|r_n\|^2 + \|r_{n+1}\|^2}{2} \, \frac{\sum\limits_{i=0}^{n} \|A^*r_i\|^{-2}}{\sum\limits_{i=0}^{n} \|r_i\|/\|Ar_i\|^2} \le C\delta \qquad (C \ge 1).$$

In [6] for iteration methods of the form

$$u_{n+1} = u_n + A^*z_n$$

the stopping index $n = n_{ME}$ is found as the first index for which

$$d_{ME}(n) \equiv \frac{(r_n + r_{n+1}, z_n)}{2\|z_n\|} \le C\delta \qquad (C \ge 1)\,.$$

In CGLS and CGME $z_n = \beta_n v_n$. The name of this rule reflects the property

$$\|u_n - u_*\| \le \|u_{n-1} - u_*\| \qquad (n = 1, 2, \ldots, n_{MA}, \ldots, n_{ME})\,. \tag{3.1}$$

For CGLS in work [1] the convergence $\|u_{n_{MA}} - u_*\| \to 0$ $(\delta \to 0)$ was proved. This also gives convergence $\|u_{n_{ME}} - u_*\| \to 0$ $(\delta \to 0)$ due to (3.1).

2) Consider now the case when there is no information about the noise level δ. Then it is principally impossible to formulate a stopping rule with convergence property $\|u_n - u_*\| \to 0$ $(\delta \to 0)$ (see [2]). However, one can find the stopping index n for example by the following Hanke-Raus rule [9]: starting with $\kappa_{-1} = 0$, $\gamma_0 = 0$, compute $\kappa_n = 1 + \sigma_n\kappa_{n-1}$, $\gamma_{n+1} = \gamma_n + \beta_n\kappa_n$ for every $n = 0,\, 1,\, 2,\, \ldots$ and find the stopping index $n = n_{HR}$ as a location of the global minimum of the function $\sqrt{\gamma_{n+1}}\|r_n\|$.

Note that in [9], for iteration methods in form $u_n = g_n(A^*A)A^*f$ with function $g_n(\lambda)$ approximating $1/\lambda$, the analogous rule was proposed: here the stopping index is a location of the global minimum of the function $\sqrt{g_{n+1}(0)}\|r_n\|$. In [9] for this stopping rule also error estimates are given.

3) Last consider the case, when noise level is known approximately: δ is given, for which it holds $\|f-f_*\|/\delta \le C$ for $\delta \to 0$ with an unknown constant C. In [5] for iterative methods of Landweber and Lardy the following stopping rule R was formulated.

Rule R. Let $0 \le s \le 1/2$. Find N as the first n for which

$$\varphi(n) \equiv \sqrt{n}\|A^*(Au_n - f)\| \le b\delta$$

with constant b large enough. Find the stopping index n_R as the location of the global minimum of the function $t(n) = n^s\|Au_n - f\|$ on the interval $[1, N]$.

For methods of Landweber and Lardy in [5] convergence $\|u_{n_R} - u_*\| \to 0$ ($\delta \to 0$) was proved and error estimates (which are quasioptimal in case $\|f - f_*\| \le \delta$) were given.

For the iterative methods CGLS and CGME the stopping index n_R may be found by an analogue of Rule R with $s \in [0, 1]$ and by replacing the function $\varphi(n)$ by function $\sqrt{\gamma_{n+1}}\|A^*(Au_n - f)\|$.

4. Numerical Experiments

We solved 10 test problems. The numerical results in tables below are given for the equation

$$(Au)(t) = \frac{1}{2\sqrt{\pi}}\int_0^1 (s-t)^{-3/2}\mathrm{e}^{-1/(4(s-t)^2)}u(s)\,\mathrm{d}s = f(t),\ 0 \le t \le 20 \quad (4.1)$$

with the solution

$$u_*(s) = \begin{cases} 3s^2/16, & \text{for } s < 2 \\ 3/4 + (s-2)(3-2), & \text{for } 2 \le s < 3 \\ 3/4\mathrm{e}^{-2(s-3)}, & \text{for } 3 \le s < 10 \\ 0, & \text{for } 10 \le s \le 20. \end{cases}$$

For the supposable noise level the values $\delta = 10^{-i}$ with $i = 1, \dots, 5$ were taken and instead of the exact data f_* randomly perturbed data were used with actual noise level $\|f - f_*\| = d\delta$ where the values of d were 1, 2, 10, 100, 1000.

The problem was discretized by the collocation method with 1024 piecewise constant basis functions on a uniform mesh and solved by the methods CGLS and CGME. In the stopping rules we used the constants $C = 1.01$, $b = 0.5$, and $s = 0.5$.

In numerical experiments we found the optimal stopping index n_* as an index n which minimizes the error $\|u_n - u_*\|$ on the interval $[1, 350]$.

In Table 1 we give for the method CGLS for case of exact noise level ($d = 1$) the stopping indexes n_*, n_D, n_{MA}, n_{ME}, n_{HR}, n_R and corresponding

Table 1. CGLS, $d = 1$, stopping indexes and errors $\|u_n - u_*\|$.

δ	n_*	n_D	n_{MA}	n_{ME}	n_{HR}	n_R	e_*	e_D	e_{MA}	e_{ME}	e_{HR}	e_R
10^{-1}	10	8	5	6	2	5	1.52	1.62	2.66	2.12	5.24	2.66
10^{-2}	22	14	11	11	9	12	0.38	0.70	1.00	1.00	1.40	0.83
10^{-3}	39	33	20	21	16	26	0.15	0.16	0.37	0.33	0.57	0.25
10^{-4}	99	55	35	35	40	50	0.09	0.10	0.14	0.14	0.12	0.11
10^{-5}	342	166	109	143	68	117	0.04	0.06	0.08	0.06	0.10	0.07

Table 2. CGLS, $d > 1$, stopping indexes.

	$d = 2$			$d = 10$			$d = 100$			$d = 1000$		
δ	n_*	n_{HR}	n_R	n_*	n_{HR}	n_R	n_*	n_{HR}	n_R	n_*	n_{HR}	n_R
10^{-1}	8	2	3	5	0	1	2	0	1	0	0	1
10^{-2}	20	7	11	10	2	5	5	0	1	2	0	1
10^{-3}	35	16	21	22	9	12	10	2	5	5	0	1
10^{-4}	61	30	39	39	16	26	22	9	12	10	2	5
10^{-5}	235	58	67	99	40	50	39	16	26	22	9	12

errors $e_* = \|u_{n_*} - u_*\|, \ldots, e_R = \|u_{n_R} - u_*\|$. For the method CGLS in the case when the actual noise level $\|f - f_*\|$ is $d > 1$ times greater than the supposed noise level δ, the stopping indexes n_*, n_{HR} and n_R are given in Table 2 and corresponding errors in Table 3 (the other rules did not stop within 350 iterations).

For the method CGME the corresponding results (without n_{MA}, e_{MA} but with n_{ME} in case $d = 2$) are given in Tables 4–6, respectively.

In case of exactly given noise level all rules tend to stop too early in all tests, with the exception of the discrepancy principle in the method CGME. In all experiments $n_{ME} \le n_D$ and $n_{MA} \le n_{ME}$, frequently $n_{HR} \le n_{ME}$. Typically we had in the method CGLS $e_D \le e_R \le e_{ME} \le e_{MA} \le e_{HR}$ and in the method CGME $e_R \le e_{ME} \le e_{HR} \le e_D$.

In case of approximately given noise level with $d \equiv \|f - f_*\|/\delta > 1$ the discrepancy principle and the monotone error rule often did not stop within 350 iterations (exception: method CGLS, $d = 2$, the monotone error rule) but the Hanke-Raus rule gave satisfactory results and rule R gave good results. However, the Hanke-Raus rule is not always applicable, since sometimes it does not stop. In the method CGLS for $d \le 10$ the rule R gave the stopping index n_R near the end of the search interval $[1, N]$, for $d = 1000$ the index n_R lies at the beginning of this interval. In the method CGME, for $d \ge 2$, in rule R the number N was not found within 350 iterations and n_R was found as the minimizer of the function $\sqrt{\gamma_{n+1}}\|A^*(Au_n - f)\|$ on the interval $[1, 350]$.

Tables 1–6 show that for our test equation (4.1) the method CGLS gave better results than the method CGME. It seems that the opposite is true for case $u_* \in \mathcal{R}(A^*)$.

Table 3. CGLS, $d > 1$, errors $\|u_n - u_*\|$.

δ	$d=2$			$d=10$			$d=100$			$d=1000$		
	n_*	e_{HR}	e_R	e_*	e_{HR}	e_R	e_*	e_{HR}	e_R	e_*	e_{HR}	e_R
10^{-1}	1.77	5.25	4.71	3.05	7.88	6.36	6.18	7.88	6.79	7.88	7.88	17.7
10^{-2}	0.62	1.85	1.04	1.52	5.24	2.66	3.05	7.88	6.36	6.18	7.88	6.79
10^{-3}	0.19	0.57	0.32	0.38	1.40	0.83	1.52	5.24	2.66	3.05	7.88	6.36
10^{-4}	0.11	0.20	0.12	0.15	0.57	0.25	0.38	1.40	0.83	1.52	5.24	2.66
10^{-5}	0.05	0.10	0.10	0.09	0.12	0.11	0.15	0.57	0.25	0.38	1.40	0.83

Table 4. CGME, $d = 1$, stopping indexes and errors $\|u_n - u_*\|$.

δ	n_*	n_D	n_{ME}	n_{HR}	n_R	e_*	e_D	e_{ME}	e_{HR}	e_R
10^{-1}	4	8	3	2	2	3.51	21.9	4.61	5.22	5.22
10^{-2}	11	14	8	7	11	1.37	3.85	1.58	1.80	1.37
10^{-3}	19	32	15	12	21	0.40	2.66	0.53	0.75	0.41
10^{-4}	36	53	32	32	36	0.14	0.46	0.15	0.15	0.14
10^{-5}	83	150	61	61	61	0.09	0.27	0.10	0.10	0.10

Table 5. CGME, $d > 1$, stopping indexes

δ	$d=2$				$d=10$			$d=100$			$d=1000$		
	n_*	n_{ME}	n_{HR}	n_R	n_*	n_{HR}	n_R	n_*	n_{HR}	n_R	n_*	n_{HR}	n_R
10^{-1}	3	7	1	2	1	0	1	0	0	1	0	0	1
10^{-2}	8	12	6	7	4	2	2	1	0	1	0	0	1
10^{-3}	15	25	12	14	11	7	11	4	2	2	1	0	1
10^{-4}	32	44	21	32	19	12	21	11	7	11	4	2	2
10^{-5}	58	111	39	55	36	32	36	19	12	21	11	7	11

Table 6. CGME, $d > 1$, errors $\|u_n - u_*\|$.

δ	$d=2$				$d=10$			$d=100$			$d=1000$		
	e_*	e_{ME}	e_{HR}	e_R	e_*	e_{HR}	e_R	e_*	e_{HR}	e_R	e_*	e_{HR}	e_R
10^{-1}	4.95	41.4	6.29	5.23	6.69	7.88	6.69	7.88	7.88	156	7.88	7.88	4254
10^{-2}	1.82	6.38	2.05	1.84	3.51	5.22	5.22	6.69	7.88	6.69	7.88	7.88	156
10^{-3}	0.60	1.97	0.76	0.63	1.37	1.80	1.37	3.51	5.22	5.22	6.69	7.88	6.69
10^{-4}	0.19	0.70	0.30	0.19	0.40	0.75	0.41	1.37	1.80	1.37	3.51	5.22	5.22
10^{-5}	0.10	0.45	0.12	0.10	0.14	0.15	0.14	0.40	0.75	0.41	1.37	1.80	1.37

References

[1] O.M. Alifanov, E.A. Artyukhin and S.V. Rumyantsev. *Extreme methods for solving ill-posed problems with applications to inverse heat transfer problems.* Begell House, New York, 1995.

[2] A.B. Bakushinskii. Remarks on the choice of regularization parameter from quasioptimality and relation tests. *Zh. Vychisl. Mat. i Mat. Fiz.*, **24**(8), 1258 – 1259, 1984. (in Russian)

[3] H.W. Engl, M. Hanke and A. Neubauer. *Regularization of inverse problems.* Kleuwer, Dordrecht, 1996.

[4] S.F. Gilyazov and N.L. Goldman. *Regularization of ill-posed problems by iteration methods.* Kluwer, Dordrecht, 2000.

[5] U. Hämarik and T. Raus. Choice of the regularization parameter in ill-posed problems with rough estimate of the noise level of data. *WSEAS Transactions on Mathematics*, **4**(2), 76 – 81, 2005.

[6] U. Hämarik and U. Tautenhahn. On the monotone error rule for parameter choice in iterative and continuous regularization methods. *BIT Numerical Mathematics*, **41**(5), 1029 – 1038, 2001.

[7] M. Hanke. *Conjugate Gradient Type Methods for Ill-Posed Problems.* Longman House, Harlow, 1995.

[8] M. Hanke. On Lanczos based methods for the regularization of discrete ill-posed problems. *BIT Numerical Mathematics*, **41**(5), 1008 – 1018, 2001.

[9] M. Hanke and T. Raus. A general heuristic for choosing the regularization parameter in ill-posed problems. *SIAM Journal on Scientific Computing*, **17**(4), 956 – 972, 1996.

[10] A.S. Nemirovski. The regularizing properties of the conjugate gradient method in ill-posed problems. *USSR Comp. Math. Math. Phys.*, **26**(2), 7 – 16, 1986.

[11] R. Plato. The method of conjugate residuals for solving the Galerkin equations associated with symmetric positive semidefinite ill-posed problems. *SIAM J. Numer. Anal.*, **35**(4), 1621 – 1645, 1998.

[12] R. Plato. The conjugate gradient method for linear ill-posed problems with operator perturbations. *Numer. Algorithms*, **20**(1), 1 – 22, 1999.

[13] R. Plato and G. Vainikko. On the fast and fully discretized solution of integral and pseudo-differential equations on smooth curves. *Calcolo*, **38**(1), 25 – 48, 2001.

[14] G. Vainikko and U. Hämarik. Projection methods and self-regularization in ill-posed problems. *Izv. Vyssh. Uchebn. Zaved. Mat.*, **10**(3), 3 – 17, 1985. (in Russian)

[15] G. Vainikko and A. Veretennikov. *Iteration Procedures in Ill-Posed Problems.* Nauka, Moscow, 1986. (in Russian)