
DISCREPANCY PRINCIPLE AND CONVERGENCE RATES IN REGULARIZATION FOR MONOTONE ILL-POSED PROBLEMS

The convergence rates of the regularized solution as well as its Galerkin approximations for nonlinear monotone ill-posed problems in Banach space are established on the basis of the choice of regularization parameter by the Morozov discrepancy principle.

На основі вибору параметра регуляризації відповідно до принципу нев’язки Морозова встановлено швидкість збіжності як регуляризованих розв’язків нелінійних моно-tonних некоректних задач у банаховому просторі, так і їх наближень Гальєркіна.

1. Introduction. Let $X$ be a real reflexive Banach space having the property: $X$ and $X^*$ are strictly convex, and weak convergence and convergence of norms of any sequence in $X$ follow from its strong convergence, where $X^*$ denotes the dual space of $X$. For the sake of simplicity the norms of $X$ and $X^*$ are denoted by the symbol $\|\cdot\|$. We write $\langle x^*, x \rangle$ or $\langle x, x^* \rangle$ instead of $x^*(x)$ for $x^* \in X^*$ and $x \in X$. Let $A$ be a monotone and continuous operator with domain $D(A) = X$ and range $D(A) \subseteq X^*$.

Our interest is to consider the operator equation

$$A(x) = f, \quad f \in D(A). \tag{1.1}$$

Without additional conditions on the structure of $A$ as strongly or uniformly monotone property, equation (1.1) is one of ill-posed problems. It means that the solutions of (1.1) depend discontinuously on the data $f$. Therefore, to find its approximative solution we have to use stable methods. A widely used and effective method is the Tikhonov regularization that consists of minimizing some functional depending on a parameter. When $A$ is nonlinear, this functional is usually non-convex. So, we cannot use the results in the theory of convex analysis to minimize the Tikhonov functional. These difficulties can be overcome for the class of problems involving monotone operators by using another version of Tikhonov regularization in form of operator equation

$$A(x) + \alpha U(x-x^0) = f_\delta, \tag{1.2}$$

where $x^0$ is some element in $X$ that plays the role of selection criterion, and $f_\delta$ are the approximations for $f$ such that $\|f_\delta - f_0\| \leq \delta$ with well-known levels $\delta \to 0$. The parameter $\alpha$ is called the parameter of regularization, and $U$ is the standard dual mapping of $X$, i. e. the mapping from $X$ onto $X^*$ that satisfies the requirement

$$\langle U(x), x \rangle = \|x\|^2, \quad \|U(x)\| = \|x\|.$$

Suppose that the following conditions hold:

$$\langle U(x) - U(y), x - y \rangle \geq \mu_U \|x - y\|^2, \quad \mu_U > 0, \quad s \geq 2, \tag{1.3}$$

$$\|U(x) - U(y)\| \leq C(R)\|x - y\|^s, \quad 0 < \delta \leq 1, \tag{1.4}$$

© NGUYEN BUONG, 2003

ISSN 0041-6053. Укр. мат. журнал, 2003, т. 55, № 7 991
where $C(R) \geq 0$, is a positive increasing function on $R = \max \{\|x\|, \|y\|\}$ (see [1]).

It is well-known (see [2, 3]) that equation (1.2) has a unique solution, henceforth denoted by $x^\delta_{\alpha}$, and if $\delta \to 0$ as $\alpha \to 0$ then the sequence $\{x^\delta_{\alpha}\}$ converges to a solution $x_0$ of (1.1) with the property

$$\|x_0 - x^0\| = \min_{x \in S_0}\|x - x^0\|,$$

where $S_0$ denotes the set of all solutions of (1.1) ($S_0 \not= \emptyset$). Moreover,

$$\|x^\delta_{\alpha} - x^0\| \leq 2\|x_0 - x^0\| + \frac{\delta}{\alpha},$$

the function $\rho(\alpha) = \alpha \|x^\delta_{\alpha} - x^0\|$ is continuous for $\alpha \in [\alpha_0, \infty)$, $\alpha_0 > 0$, and

$$\lim_{\alpha \to \infty} \rho(\alpha) = \|x^0 - f_\delta\|.$$ For each fixed $\delta > 0$ the value $\alpha$, that satisfies the condition $\delta / \alpha \to 0$ as $\delta \to 0$, can be chosen by the Morozov discrepancy principle formulated as follows.

Assume that $\|A(x^0) - f_\delta\| > K\delta^p$, $0 < p < 1$, $K > 1$. For $0 < \delta < \delta_0 < 1$, there exists a value $\delta_0$ such that $\rho(\delta_0) = K\delta^p$, $\delta_0 \to 0$ as $\delta \to 0$ and $x^\delta_{\delta_0} \to x_0$, $\delta / \delta_0 \to 0$ as $\delta \to 0$ (see [3, 4]).

The solution $x^\delta_{\alpha}$ can be approximated by the solutions of the finite-dimensional problems

$$A^\alpha U^\alpha(x - x^{0, u}) = f_\delta^\alpha,$$ (1.5)

where $A^\alpha = P_n^\alpha A P_n$, $U^\alpha(x) = P_n^\alpha U P_n(x)$, $f_\delta^\alpha = P_n^\alpha f_\delta$, $x^{0, u} = P_n x^0$, $P_n$ denotes a (linear) projection from $X$ onto its subspace $X_n$, $P_n^\ast$ is the adjoint of $P_n$ and $X_n \subset X_{n+1}$ $\forall n$, $P_n x \to x$ $\forall x \in X$ ($\|P_n\| = 1$).

For each $\alpha > 0$ equation (1.5) has a unique solution $x_{\alpha, n}$, and the sequence $\{x_{\alpha, n}\}$ converges to $x^\delta_{\alpha}$ as $n \to +\infty$ (see [5]). It is very important for computation to know convergence rates of the sequence $\{x_{\alpha, n}\}$, whether

$$\lim_{n \to \infty} x_{\alpha, n} = x_0$$

and convergence rates of the sequence $\{x_{\alpha, n}\}$.

For the linear ill-posed problems these questions are completely studied when the values $\alpha = \alpha(\delta)$ is chosen arbitrarily or by the Morozov discrepancy principles (see [6 – 11]). For nonlinear case these questions were studied in [12 – 16]. In [17] and [18] the convergence rates of the regularized solutions of the ill-posed equation involving nonlinear monotone operator in Banach space are investigated when the regularization parameter is chosen arbitrarily such that $\alpha \sim \delta^p$. $0 < p < 1$.

In this paper, by using the discrepancy principle for parameter choice we obtain the estimate for convergence rates of the regularized solution as well as its Galerkin approximations. In particular, the obtained results here are guaranteed under weaker conditions than in [17, 18].

As usually, the symbols $\rightarrow$ and $\Rightarrow$ denote weak convergence and convergence in norm, respectively, and the notation $a \sim b$ means that $a = O(b)$ and $b = O(a)$.

In the following section we suppose that all above conditions are satisfied.
2. Main results.

Assumption 2.1. There exists a constant $\tau > 0$ such that for $y$ in some neighbourhood of $S_0$ and $x \in S_0$ the following relation is true:

$$\|A(y) - A(x) - A'(x)(y - x)\| \leq \tau \|y - x\| \|A'(x)(y - x)\|.$$

This condition is illustrated by concrete problems in [19].

First, we prove a result about convergence rates for $\{x_\alpha^\delta\}$.

Theorem 2.1. Assume that the following conditions hold:

(i) $A$ is Fréchet differentiable at some neighbourhood of $S_0$ with Assumption 2.1 for $x = x_0$:

(ii) there exists an element $z \in X$ such that

$$A'(x_0)^*z = U(x_0 - x^0);$$

(iii) the parameter $\alpha = \alpha(\delta)$ is chosen by the discrepancy principle.

Then

$$\|x_\alpha^\delta - x_0\| = O(\delta^\theta), \quad \theta = \min \left\{ \frac{1 - p}{s - 1}, \frac{p}{s} \right\}.$$

Proof. From (1.1) – (1.3) and condition (ii) of the Theorem it follows

$$m_u \|x_\alpha^\delta - x_0\|^p \leq (U(x_\alpha^\delta - x_0) - U(x_0 - x^0), x_\alpha^\delta - x_0) \leq$$

$$\leq \frac{1}{\alpha} \langle f_\alpha - A(x_\alpha^\delta), x_\alpha^\delta - x_0 \rangle + (U(x_0 - x^0), x_0 - x_\alpha^\delta) \leq$$

$$\leq \frac{\delta}{\alpha} \|x_\alpha^\delta - x_0\| + \|z\| \|A'(x_0)(x_0 - x_\alpha^\delta)\|.$$

(2.1)

On the other hand, from Assumption 2.1 it implies that

$$\|A'(x_0)(x_0 - x_\alpha^\delta)\| \leq \|A(x_\alpha^\delta) - f_\delta\| + \tau \|x_\alpha^\delta - x_0\| \|A'(x_0)(x_\alpha^\delta - x_0)\| \leq$$

$$\leq \|A(x_\alpha^\delta) - f_\delta\| + \delta + \tau \|x_\alpha^\delta - x_0\| \|A'(x_0)(x_\alpha^\delta - x_0)\|.$$

If $\alpha = \alpha(\delta)$ is chosen by the discrepancy principle, then for sufficiently small $\delta$ we have $\tau \|x_\alpha^\delta - x_0\| < 1/2$. Hence,

$$\|A(x_\alpha^\delta) - f_\delta\| = \alpha \|x_\alpha^\delta - x^0\| = \rho(\alpha) = K\delta^\rho,$$

$$\|A'(x_0)(x_\alpha^\delta - x_0)\| \leq 2\|A(x_\alpha^\delta) - f_\delta\| + \delta \leq$$

$$\leq 2(K\delta^\rho + \delta) = 2(K + 1)\delta^\rho$$

and

$$\frac{\delta}{\alpha} \leq 2 \|x_0 - x_0\| \|K^{-1}\| \delta^{1 - p}$$

(see [3] or [4]). Consequently, from (2.1) we obtain

$$m_u \|x_\alpha^\delta - x_0\|^p \leq 2 \|x_0 - x_0\| \|K^{-1}\| \delta^{1 - p} \|x_\alpha^\delta - x_0\| +$$

$$+ 2 \|z\| \|K + 1\| \delta^\rho.$$

Using the implication

ISSN 0041-6053. Укра̀м. мат. журÒ., 2003, t. 55, № 7
\(a, b, c \geq 0, \quad p > q, \quad a^p \leq ba^q + c \Rightarrow a^p = O(b^{n(p-q)} + c)\)

we obtain

\[\|x_{\alpha}^\delta - x_0\| = O(\delta^\theta), \quad \theta = \min\left\{\frac{1-p}{s-1}, \frac{p}{s}\right\}.\]

The Theorem 2.1 is proved.

As spoken above, for numerical implementation one should consider the finite-dimensional approximation for \(x_{\alpha}^\delta\) by the solutions of (1.5). Therefore, the problem of selecting the value \(\alpha\) by the discrepancy principle is necessary to consider in relation with finite-dimensional approximations of the space \(X\). First of all, we see that if \(A\) is coercive or \(\langle A(x) - f, x \rangle > 0, \|x\| > r, r > 0\), i.e., (1.1) has solution, then \(A_n\) possesses the same properties, i.e., equation

\[A^n(x) = f^n, \quad f^n = P_n^* f,\] (2.2)

has solution, too. Thus equation (1.5) can be viewed as regularized equation for (2.2). It is reasonable to remark that the solution \(x_n\) of (2.2) does not always converge to a solution of (1.1) as \(n \to \infty\) (see [20]). On the other hand, it is easy to see that

\[\|A^n(x_0^n) - f_\delta^n\| = \|P_n^*(A^n(x_0^n) - f_\delta^n)\| \to \|A(x) - f_\delta\|\]

as \(n \to \infty\). Because of the last inequality and \(\|A(x_0^n) - f_\delta\| > K\delta^p\) we have \(\|A^n(x_0^n) - f_\delta^n\| > K\delta^p\) for sufficiently large \(n\).

On the other hand,

\[\|x_{\alpha,n}^\delta - x_0^n\| \leq \|x_0^n - x_0^n\| + \frac{\delta}{\alpha} \leq 2\|x_0^n - x_0^n\| + \frac{\delta}{\alpha},\]

\[\|f_\delta^n - f^n\| = \|P_n^*(f_\delta - f)\| \leq \delta, \quad x_0^n = P_nx_0^n.\]

For sufficiently small \(\alpha\) such that

\[2\alpha\|x_0^n - x_0^n\| \leq (K-1)\delta^p,\]

we have the inequality

\[\alpha\|x_{\alpha,n}^\delta - x_0^n\| \leq 2\alpha\|x_0^n - x_0^n\| + \delta \leq K\delta^p.\]

Therefore, the value \(\alpha\) can be chosen by \(\rho_n(\alpha) = K\delta^p\) and

\[\alpha > \frac{(K-1)\|x_0^n - x_0^n\|}{2} + c\gamma_n^p,\]

where

\[\rho_n(\alpha) = \|A^n(x_{\alpha,n}^\delta) - f_\delta^n\|,\]

\[\gamma_n^p = \max\{\gamma_n(x_0^n), \gamma_n(x_0^n), \gamma_n(x), \gamma_n^*(f)\},\]

\[\gamma_n(x) = \|(I - P_n)x\|, \quad \gamma_n^*(f) = \|(I - P_n^*)x\|\]

and \(c\) is a positive constant. This way in finding the parameter \(\alpha\) requires to solve the equation \(\|A^n(x_{\alpha,n}^\delta) - f_\delta^n\| = K\delta^p\). This is a very complex work. So, we should use its slightly modification as follows.

The rule:

choose \(\alpha \geq (c_1\delta + c_2\gamma_n)^p, \quad c_1 > 1, \quad 0 < p < 1\) such that

ISSN 0041-6053. Укр. мат. журн., 2003, т. 55, № 7
\[ \| A^n(\delta) \alpha_n - f_n^\delta \| \geq K \delta^n. \]
\[ \| A^n(\delta) \alpha_n - f_n^\delta \| \leq K_1 \delta^n, \quad K_1 \geq K. \]

As in [3], we can verify that \( \overline{\alpha}(\gamma_n) \rightarrow 0 \), \( (\delta + \gamma_n)/\overline{\alpha}(\gamma_n) \rightarrow 0 \) as \( \delta \rightarrow 0 \)
and \( n \rightarrow \infty \). Let \( x_{\alpha,n}^\delta \) be the solution of (1.5) with \( \alpha = \overline{\alpha} \). The convergence and rate of convergence \( \{ x_{\alpha,n}^\delta \} \) to \( x_0 \) as \( \delta \rightarrow 0 \) and \( n \rightarrow \infty \) are determined by the following theorems.

**Theorem 2.2.** If \( \delta/\alpha, \gamma_n(x)/\alpha \rightarrow 0 \) with \( x \in S_0 \) as \( \delta, \alpha \rightarrow 0 \) and \( n \rightarrow \infty \), then the sequence \( \{ x_{\alpha,n}^\delta \} \) converges to \( x_0 \).

**Proof.** For \( x \in S_0 \), \( x^n = P_n x \), from (1.3) and (1.5) it follows
\[
\begin{align*}
m_U \| x_{\alpha,n}^\delta - x^n \| \leq & \langle U(x_{\alpha,n}^\delta - x^{0,n}, x_{\alpha,n}^\delta - x^n) \rangle \\
\leq & \langle U^n(x_{\alpha,n}^\delta - x^{0,n}, x_{\alpha,n}^\delta - x^n) + A(x^n - x^{0,n}, x^n - x_{\alpha,n}^\delta) \rangle \\
& + \langle U(x^n - x^{0,n}, x^n - x_{\alpha,n}^\delta) \rangle.
\end{align*}
\]

On the other hand,
\[
\| A(x^n) - A(x) \| \leq \| A'(x)(I - P_n)x \| + \tau \| (I - P_n)x \| \| A'(x)(I - P_n)x \| \leq C_0 \gamma_n(x)(\tau \gamma_n(x) + 1),
\]
where \( C_0 \) is some positive constant depending only on \( x \). Therefore, from (2.3) we have got
\[
\begin{align*}
m_U \| x_{\alpha,n}^\delta - x^n \| \leq & \frac{\delta + C_0 \gamma_n(x)(\tau \gamma_n(x) + 1)}{\alpha} \| x_{\alpha,n}^\delta - x^n \| \\
& + \langle U(x^n - x^{0,n}, x^n - x_{\alpha,n}^\delta) \rangle.
\end{align*}
\]

Since \( \delta/\alpha, \gamma_n(x)/\alpha \rightarrow 0 \) as \( \delta, \alpha \rightarrow 0 \) and \( n \rightarrow \infty \), the sequence \( \{ x_{\alpha,n}^\delta \} \) is bounded. Without loss of generality, let \( x_{\alpha,n}^\delta \rightarrow x_1 \) as \( \delta, \alpha \rightarrow 0 \) and \( n \rightarrow +\infty \). We write the monotone property for \( A^n \) as follows
\[
\langle A^n(x^n) - A^n(x_{\alpha,n}^\delta), x^n - x_{\alpha,n}^\delta \rangle \geq 0 \quad \forall x \in X, \quad x^n = P_n x.
\]

Because \( P_n^*P_n = P_n^* \), the last inequality is transformed into the form
\[
\langle A(x^n) - A^n(x_{\alpha,n}^\delta), x^n - x_{\alpha,n}^\delta \rangle \geq 0.
\]

Hence,
\[
\langle A(x^n) - f_\delta, x^n - x_{\alpha,n}^\delta \rangle + \alpha \| x_{\alpha,n}^\delta - x^{0,n} \| \| x^n - x_{\alpha,n}^\delta \| \geq 0.
\]

Passing to the limit as \( \delta, \alpha \rightarrow 0 \) and \( n \rightarrow +\infty \) in this inequality we obtain
\[
\langle A(x) - f, x - x_1 \rangle \geq 0 \quad \forall x \in X.
\]

By Minty lemma (see [20, p. 257]) \( x_1 \in S_0 \).

On the other hand, from (2.4) we also obtain \( \langle U(x - x^0), x - x_1 \rangle \geq 0 \quad \forall x \in S_0 \).

Because of the convex and closed property of \( S_0 \) we have \( \langle U(tx_1 + (1 - t)x - x^0), x_1 \rangle \geq 0 \quad \forall x \in S_0. \)

**ISSN 0041-6053. Укр. мат. журн., 2003, т. 55, № 7**
\( x - x_1 \geq 0 \ \forall \ x \in S_0, \ t \in (0, 1) \). Passing to the limit as \( t \to 1 \) in this inequality we get \( \langle U(x_1 - x^0), x - x_1 \rangle \geq 0 \ \forall \ x \in S_0 \). Consequently, \( \| x_1 - x^0 \| \leq \| x - x^0 \| \ \forall \ x \in S_0 \). Since \( S_0 \) is a convex and closed subset of a strictly convex Banach space \( X \), then \( x_1 = x^0 \). Therefore, the entire sequence \( \{ x_{\alpha, n}^\delta \} \) converges weakly to \( x_0 \).

Replacing \( x^n \) by \( x_{\alpha, n}^\delta \) in (2.4) we can see that the sequence \( \{ x_{\alpha, n}^\delta \} \) converges strongly to \( x_0 \) as \( \delta, \alpha \to 0 \) and \( n \to +\infty \). Theorem is proved.

**Theorem 2.3.** Suppose that the following conditions hold:

(i) \( A \) is a Fréchet differentiable at some neighbourhood of \( S_0 \) with Assumption 2.1 for \( x_1 = x_0 \);

(ii) there exists an element \( z \in X \) such that

\[ A'(x_0)^* z = U(x_0 - x^0) \; ; \]

(iii) the parameter \( \bar{\alpha} = \alpha(\delta, \gamma_n) \) is chosen by the rule.

Then

\[ \| x_{\alpha, n}^\delta - x_0 \| = O((\delta + \gamma_n)^{(1-p)/2(x-1)} + \gamma_n^{\delta/(x-1)} + (\delta^p + \gamma_n)^{1/x}) \; . \]

**Proof.** Replacing \( x \) in (2.4) by \( x_0 \) we obtain

\[ m_U \| x_{\alpha, n}^\delta - x_0 \| \leq \]

\[ \leq \frac{\delta + C_0 \gamma_n (\tau y_n + 1)}{\alpha} \| x_{\alpha, n}^\delta - x_0 \| + \langle U(x_0 - x^0), x_{\alpha, n}^\delta - x_{\alpha, n} \rangle + \]

\[ + \langle U(x_0 - x_0^n), U(x_0 - x^0), x_{\alpha, n}^\delta - x_{\alpha, n} \rangle \; . \]

(2.5)

From (1.4) and Assumption 2.1 it follows that

\[ \| \langle U(x_0^n - x_0^n), U(x_0 - x^0), x_{\alpha, n}^\delta - x_{\alpha, n} \rangle \| \leq C(\bar{R}) 2^\delta \gamma_0 \| x_{\alpha, n}^\delta - x_0 \| \; , \]

where \( \bar{R} \geq \| x_0 - x^0 \| \), and

\[ \| \langle U(x_0 - x^0), x_0^n - x_{\alpha, n}^\delta \rangle \| \leq \| \langle U(x_0 - x^0), x_0^n - x_0 \rangle \| + \]

\[ + \gamma_n \| A'(x_0)(x_{\alpha, n}^\delta - x_0) \| + \| \langle z^n, A'(x_0)(x_{\alpha, n}^\delta - x_0) \rangle \| \leq \]

\[ \leq (\bar{R} + \| A'(x_0)(x_{\alpha, n}^\delta - x_0) \|) \gamma_n + \| \langle z^n, A'(x_0)(x_0 - x_{\alpha, n}^\delta) \rangle \| \; . \]

On the other hand,

\[ \| \langle z^n, A'(x_0)(x_0 - x_{\alpha, n}^\delta) \rangle \| \leq \]

\[ \leq \| \langle z^n, f - A(x_{\alpha, n}^\delta) \rangle \| + \| \langle z^n, A(x_{\alpha, n}^\delta) - A(x_0) - A'(x_0)(x_{\alpha, n}^\delta - x_0) \rangle \| \leq \]

\[ \leq \| z \| \| (A'(x_{\alpha, n}^\delta) - f_0^\delta) \| + \delta \| x_{\alpha, n}^\delta - x_0 \| \| \langle z^n, A'(x_0)(x_0 - x_{\alpha, n}^\delta) \rangle \| \; . \]

Since \( \| x_{\alpha, n}^\delta - x_0 \| \to 0 \), we can write \( \tau \| x_{\alpha, n}^\delta - x_0 \| < 1/2 \) for sufficiently small \( \alpha \), \( \beta \) and large \( n \). So, when \( \alpha \) is chosen by the rule we have \( \| A'(x_{\alpha, n}^\delta) - f_0^\delta \| \leq K_1 \delta^p \). Consequently,

\[ \| \langle z^n, A'(x_0)(x_0 - x_{\alpha, n}^\delta) \rangle \| \leq 2 \| z \| \| (A'(x_{\alpha, n}^\delta) - f_0^\delta) \| + \delta \leq 2 \| z \| (K_1 + 1) \delta^p \; . \]

Therefore,
\[ \left| \langle U(x_0 - x^0), x_n - x^\delta_{\alpha,n} \rangle \right| \leq \left( \tilde{R} + \left\| A'(x_0)(x^\delta_{\alpha,n} - x_0) \right\| \right) \gamma_n + 2 \| z \| (K_1 + 1) \delta^p. \]

Finally, inequality (2.5) has the form
\[ m_U \left\| x^\delta_{\alpha,n} - x_0^0 \right\|^x \leq \leq C_1 \left( (\delta + \gamma_n)^{1-p} + \gamma_0^0 \right) \left\| x^\delta_{\alpha,n} - x_0^0 \right\| + C_2 \delta^p + C_3 \gamma_n, \quad C_i > 0. \]

Thus,
\[ \left\| x^\delta_{\alpha,n} - x_0^0 \right\| \leq O \left( (\delta + \gamma_n)^{1-p}/(x-1) + \gamma_0^0/(x-1) + (\delta^p + \gamma_0^0)^{1/x} \right). \]

Hence,
\[ \left\| x^\delta_{\alpha,n} - x_0^0 \right\| \leq O \left( (\delta + \gamma_n)^{1-p}/(x-1) + \gamma_0^0/(x-1) + (\delta^p + \gamma_0^0)^{1/x} \right). \]

Theorem is proved.


Received 21.01.2002