ЖУРНАЛ ВЫЧИСЛИТЕЛЬНОЙ МАТЕМАТИКИ И МАТЕМАТИЧЕСКОЙ ФИЗИКИ, 2023, том 63, № 5, с. 840–855

МАТЕМАТИЧЕСКАЯ ФИЗИКА

УДК 517.988

МЕТОД КВАЗИРЕШЕНИЙ И ПРОБЛЕМА ГЛОБАЛЬНОЙ МИНИМИЗАЦИИ ФУНКЦИОНАЛА НЕВЯЗКИ УСЛОВНО КОРРЕКТНЫХ ОБРАТНЫХ ЗАДАЧ

© 2023 г. М. Ю. Кокурин[1,*]

[1] *424001 Йошкар-Ола, пл. Ленина, 1, Марийский государственный университет, Россия*

**e-mail: kokurinm@yandex.ru*

Поступила в редакцию 20.11.2021 г.
Переработанный вариант 12.11.2022 г.
Принята к публикации 02.02.2023 г.

Рассматривается класс условно корректных задач, характеризуемый гёльдеровой оценкой условной устойчивости на выпуклом компакте в гильбертовом пространстве. Оператор прямой задачи и правая часть уравнения заданы с погрешностями, близость производных точного и возмущенного операторов не предполагается. Исследуются свойства выпуклости и одноэкстремальности функционала невязки метода квазирешений. Для этого функционала устанавливается, что каждая его стационарная точка на множестве условной корректности, не слишком далекая от искомого решения исходной обратной задачи, лежит в малой окрестности решения. Даны оценки диаметра указанной окрестности в терминах погрешностей входных данных. Показано, что эта окрестность является аттрактором итераций метода проекции градиента, и получены оценки скорости сходимости итераций к аттрактору. Устанавливается необходимость используемой оценки условной устойчивости для существования итерационных процессов с указанными свойствами. Библ. 16.

Ключевые слова: обратная задача, условно корректная задача, метод квазирешений, глобальная оптимизация, оценка точности, эффект кластеризации, метод проекции градиента, аттрактор, скорость сходимости.

DOI: 10.31857/S0044466923050137, **EDN:** PKJKKS

1. ПОСТАНОВКА ЗАДАЧИ

В работе рассматриваются нелинейные обратные задачи, описываемые операторными уравнениями

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

где $F : H_1 \to H_2$ – оператор прямой задачи, предполагаемый дифференцируемым по Фреше на выпуклом компактном множестве $D \subset H_1$; H_1, H_2 – вещественные гильбертовы пространства. В приложениях множество D определяет априорные ограничения на искомый элемент $x \in H_1$. Этот элемент описывает подлежащий определению набор параметров исследуемой модели. Элемент $f \in H_2$ описывает результаты наблюдения и выполняет роль входных данных для обратной задачи реконструкции параметров модели. Всюду далее считаем, что оператор F инъективен на множестве D и $f \in F(D)$. Обозначим через $x^* = F^{-1}(f)$ искомое решение задачи (1.1). Ниже будем предполагать, что производная F' удовлетворяет условию Липшица

$$\|F'(x) - F'(y)\|_{L(H_1,H_2)} \le L\|x - y\|_{H_1} \quad \forall x, y \in D. \tag{1.2}$$

Из (1.2) следует, что

$$\|F(x) - F(y)\|_{H_2} \le M\|x - y\|_{H_1}, \quad \|F'(x)\|_{L(H_1,H_2)} \le M \quad \forall x, y \in D. \tag{1.3}$$

В (1.3) можно положить $M = \|F'(\bar{x})\|_{L(H_1,H_2)} + Ld_0$, $\bar{x} \in D$. Здесь и далее $d_0 = \operatorname{diam}_{H_1}(D)$,

$$\operatorname{diam}_X(\Omega) \triangleq \sup\{\|x - y\|_X : x, y \in \Omega\}, \quad \Omega \subset X.$$

Через $\|\cdot\|_X$ обозначаем норму в гильбертовом или банаховом пространстве X, $(\cdot,\cdot)_X$ есть скалярное произведение в гильбертовом пространстве X. Следствием (1.2) и интегральной теоремы о конечных приращениях (см. [1, § 32]) является представление

$$F'(x)(x-y) = F(x) - F(y) + G(x,y),$$
$$\|G(x,y)\|_{H_2} \le \frac{1}{2}L\|x-y\|^2_{H_1}, \quad x,y \in D. \tag{1.4}$$

Ввиду неизбежных погрешностей измерения, элемент f и оператор F, как правило, бывают заданы приближенно, так что вместо них доступны аппроксимации $\tilde{f} \in H_2$ и $\tilde{F}: H_1 \to H_2$. Конечномерные аппроксимации F, используемые при построении численно реализуемых процедур решения (1.1), в ряде случаев также удобно рассматривать в качестве приближенных операторов $\tilde{F}$. Типичные требования к качеству аппроксимации сводятся к тому, что (см. [2, гл. 2, § 4])

$$\|\tilde{f} - f\|_{H_2} \le \delta, \tag{1.5}$$

оператор $\tilde{F}$ дифференцируем по Фреше и, кроме того,

$$\|\tilde{F}(x) - F(x)\|_{H_2} \le h, \tag{1.6}$$

$$\|\tilde{F}'(x) - F'(x)\|_{L(H_1,H_2)} \le h \quad \forall x \in D. \tag{1.7}$$

Величины δ, h в (1.5)–(1.7) характеризуют уровень погрешности в задании элемента f и оператора F соответственно.

На практике условие (1.7), предполагающее равномерную близость производных исходного и возмущенного операторов, часто оказывается излишне ограничительным. Например, при исследовании проекционных схем аппроксимации уравнения (1.1) вводятся операторы $\tilde{F}(x) = F(\mathcal{P}_N x)$, где $\{\mathcal{P}_N\}$ есть семейство ортопроекторов из H_1 на конечномерные подпространства $\mathcal{H}_N$:

$$\mathcal{H}_N \subset \mathcal{H}_{N+1}, \quad N = 1,2,\ldots; \quad \overline{\bigcup_{N=1}^{\infty}\mathcal{H}_N} = \mathcal{H}_1.$$

В этом случае условие (1.6) имеет место с $h = M\gamma_N(D)$, где

$$\gamma_N(D) = \max\{\|(E - \mathcal{P}_N)x\|_{H_1} : x \in D\}.$$

При этом $\lim\limits_{N\to\infty}\gamma_N(D) = 0$ в силу компактности множества D. Однако, если оператор F не является вполне непрерывным, то условие (1.7), вообще говоря, выполняется лишь при $h \ge h_1 > 0$ с фиксированной постоянной h_1, в чем убеждает пример $F(x) = x$. Эти соображения мотивируют введение следующего ослабленного по сравнению с (1.7) требования:

$$\|\tilde{F}'(x)\|_{L(H_1,H_2)} \le M_0, \quad \|(\tilde{F}'(x) - F'(x))(x - x^*)\|_{H_2} \le \mu \quad \forall x \in D. \tag{1.8}$$

Здесь μ – мера возмущения производной оператора F, предполагаемая достаточно малой. Нетрудно видеть, что условие (1.8) является следствием (1.7), при этом можно принять $\mu = d_0 h$. Пусть теперь $\tilde{F}(x) = F(\mathcal{P}_N x)$, и неравенства (1.2), (1.3) имеют место для всех $x, y \in O_{R_0}(0) \supset D$:

$$O_R(x) \triangleq \{y \in H_1 : \|x - y\|_{H_1} \le R\}.$$

Тогда (1.8) выполнено с $M_0 = M$ и $\mu = \mu_N = (Ld_0 + 2M)\gamma_N(D)$, при этом $\lim\limits_{N\to\infty}\mu_N = 0$. В дальнейшем будем считать, что $h + \mu + \delta \le \Delta_0$ для некоторого фиксированного Δ_0.

Характерное свойство многих прикладных обратных задач, в частности, обратных задач математической физики (см. [3–6]) состоит в том, что обратный оператор F^{-1} как отображение из H_2 в H_1 не является непрерывным в точках из $F(D)$ в том смысле, что из сходимости

$$\lim_{n\to\infty}\|f_n - f\|_{H_2} = 0, \quad f_n \in H_2, \quad f \in F(D),$$

не следует сходимость элементов $F^{-1}(f_n)$, если они определены, к $F^{-1}(f)$ в метрике H_1. В этом случае уравнение (1.1) является некорректной по Адамару задачей. Широко распространенным подклассом некорректных обратных задач (1.1) является класс условно корректных (корректных по А.Н. Тихонову) на D задач (см. [3], [4], [7], [8]). Условная корректность задачи (1.1) на множестве D по определению означает, что оператор F^{-1} относительно непрерывен на $F(D)$, т.е. равенство

$$\lim_{n\to\infty}\|f_n - f\|_{H_2} = 0, \quad f_n, f \in F(D),$$

влечет $\lim\limits_{n\to\infty}\left\|F^{-1}(f_n) - F^{-1}(f)\right\|_{H_1} = 0$. Здесь требуется, чтобы аппроксимирующие $f \in F(D)$ элементы f_n также лежали в $F(D)$. В наших предположениях (1.1) условно корректна ввиду компактности D.

Классический подход к построению устойчивых процедур аппроксимации решений некорректных задач связан с понятием регуляризующего оператора (алгоритма). Ниже обратимся к методу квазирешений В.К. Иванова (см. [3], [9]). Регуляризующий оператор, отвечающий методу квазирешений, имеет вид

$$R(\tilde F, \tilde f) = \arg\min_{x\in D} \tilde J(x), \quad \tilde J(x) = \frac{1}{2}\left\|\tilde F(x) - \tilde f\right\|^2_{H_2}. \tag{1.9}$$

Таким образом, в качестве приближения к искомому решению x^* выбирается элемент из D, доставляющий глобальный минимум функционалу $\tilde J$ на множестве D. Ввиду непрерывности оператора $\tilde F$ и компактности D, множество $\tilde X^*$ решений задачи (1.9) не пусто. При этом каждый элемент $x = \tilde x^* \in \tilde X^*$ удовлетворяет необходимому условию минимума (см. [10, гл. 1, § 2])

$$(\tilde J'(\tilde x), \tilde x - x)_{H_1} \le 0 \quad \forall x \in D. \tag{1.10}$$

Градиент функционала $\tilde J$ имеет вид

$$\tilde J'(x) = \tilde F'^*(x)(\tilde F(x) - \tilde f). \tag{1.11}$$

Точки $\tilde x \in D$, удовлетворяющие условию (1.10), будем называть стационарными точками функционала $\tilde J$ на D.

Схема (1.9), как и близкая к ней схема А.Н. Тихонова (см. [9]), труднореализуема численно в случае произвольного гладкого оператора $\tilde F$. Причина состоит в том, что задача (1.9), вообще говоря, многоэкстремальная и, следовательно, трудноразрешима стандартными итерационными методами. Это же относится и к конечномерным аппроксимациям (1.9), используемым при конструировании численно реализуемых методов решения (1.1). Непосредственное применение к указанным задачам классических итерационных методов минимизации в лучшем случае приводит лишь к некоторой стационарной точке, близость которой к глобальному минимуму в (1.9) априори не гарантируется. Отсутствие свойств выпуклости минимизируемого функционала в (1.9) создает дополнительные трудности при исследовании поведения стандартных процедур минимизации (см., например, [10, гл. 1, § 4]). При сделанных выше предположениях большинство таких процедур гарантируют лишь сходимость подходящих подпоследовательностей к соответствующим стационарным точкам. Тем не менее ниже будет показано, что для широкого класса задач за счет обозримых дополнительных условий на их характеристики может быть обеспечена определенная регулярность итераций метода проекции градиента. Нам понадобится требование равномерной условной корректности задачи (1.1).

Условие 1. Имеет место степенная оценка условной устойчивости задачи (1.1) с показателем $p \in [1,2]$:

$$\|F(x) - F(y)\|_{H_2} \ge m\|x - y\|^p_{H_1} \quad \forall x, y \in D; \quad m > 0. \tag{1.12}$$

Условие 1 выполняется для многих обратных задач математической физики (см., например, [5], [6]).

Далее будет установлено, что любая не слишком удаленная от x^* стационарная точка задачи (1.9) на самом деле лежит в малой окрестности искомого решения x^*. Тем самым, малая окрестность x^* содержит кластер стационарных точек (1.9), удаленный от других стационарных точек этой задачи. Будет дана оценка диаметра этой окрестности в терминах уровней погрешности h, μ, δ. Кроме того покажем, что указанная окрестность является аттрактором итераций метода проекции градиента и получим оценки скорости сходимости итерационной последовательности к этому аттрактору в зависимости от значений p.

План работы следующий. В разд. 2 в предположении $h = 0$ исследуются свойства локальной сильной выпуклости и одноэкстремальности функционала невязки (1.9) в зависимости от значений показателя p из условия 1. В разд. 3 приводятся примеры образования кластеров стационарных точек $\tilde{J}$ на D в окрестности решения при возмущениях оператора, удовлетворяющих условиям (1.7) и (1.8). Раздел 4 посвящен доказательству основных результатов о локализации кластеров стационарных точек и сходимости к ним итераций метода проекции градиента. В разд. 5 устанавливается необходимость условия 1 для существования итерационных процессов с указанными свойствами.

2. СЛУЧАЙ ТОЧНО ЗАДАННОГО ОПЕРАТОРА

Изучим свойства локальной выпуклости и одноэкстремальности функционала задачи (1.9). Содержательные результаты здесь удается получить лишь в случае точно заданного оператора, т.е. при $\tilde{F} = F$. В большинстве случаев приходится также предполагать, что $\tilde{f} = f$. Рассмотрим отдельно два возможных случая, отличающихся значением показателя p в условии (1.12).

1) Пусть $p = 1$. Убедимся, что при $h = 0$ и достаточно малом δ функционал $\tilde{J}$ является сильно выпуклым в окрестности точного решения x^* уравнения (1.1). Воспользуемся известным критерием сильной выпуклости (см. [10, гл. 1, § 2]). Для произвольных $x, y \in D$ с учетом (1.4) имеем

$$\begin{gathered}(\tilde{J}'(x) - \tilde{J}'(y), x - y)_{H_1} = (F'^*(x)(F(x) - \tilde{f}) - F'^*(y)(F(y) - \tilde{f}), x - y)_{H_1} = \\ = (F(x) - F(y), F'(x)(x - y))_{H_2} + (F(y) - \tilde{f}, (F'(x) - F'(y))(x - y))_{H_2} = \\ = (F(x) - F(y), F(x) - F(y) + G(x, y))_{H_2} + (F(y) - \tilde{f}, (F'(x) - F'(y))(x - y))_{H_2} \ge \\ \ge \|F(x) - F(y)\|_{H_2}^2 - \frac{1}{2}L\|F(x) - F(y)\|_{H_2}\|x - y\|_{H_1}^2 - L\|F(y) - \tilde{f}\|_{H_2}\|x - y\|_{H_1}^2.\end{gathered}$$

Отсюда с использованием (1.12) при $p = 1$ и (1.3), (1.5) получаем

$$\begin{gathered}(\tilde{J}'(x) - \tilde{J}'(y), x - y)_{H_1} \ge \|F(x) - F(y)\|_{H_2}^2 - \frac{L}{2m^2}\|F(x) - F(y)\|_{H_2}^3 - L(\|F(y) - f\|_{H_2} + \delta) \times \\ \times \|x - y\|_{H_1}^2 \ge \|F(x) - F(y)\|_{H_2}^2\left(1 - \frac{LM}{2m^2}\|x - y\|_{H_1}\right) - L(M\|y - x^*\|_{H_1} + \delta)\|x - y\|_{H_1}^2.\end{gathered} \tag{2.1}$$

Пусть теперь $x, y \in D \cap O_R(x^*)$. Тогда из (2.1) следует

$$(\tilde{J}'(x) - \tilde{J}'(y), x - y)_{H_1} \ge \|F(x) - F(y)\|_{H_2}^2\left(1 - \frac{LMR}{m^2}\right) - L(MR + \delta)\|x - y\|_{H_1}^2. \tag{2.2}$$

Предположим, что

$$2LMR + L\delta \le m^2 - \nu, \quad \nu > 0. \tag{2.3}$$

Используя (1.12), из (2.2) и (2.3) получаем

$$(\tilde{J}'(x) - \tilde{J}'(y), x - y)_{H_1} \ge (m^2 - 2LMR - L\delta)\|x - y\|_{H_1}^2 \ge \nu\|x - y\|_{H_1}^2.$$

Тем самым доказана следующая теорема, устанавливающая локальную сильную выпуклость $\tilde{J}$ в окрестности решения x^*.

Теорема 1. *Пусть* $h = 0$*, выполняются условие* 1 *при* $p = 1$ *и условия* (1.5), (2.3). *Тогда функционал* $\tilde{J}$ *является сильно выпуклым на* $D \cap O_R(x^*)$.

Следствие 1. В условиях теоремы 1 функционал $\tilde{J}$ имеет не более одной стационарной точки в $D \cap O_R(x^*)$.

Условие (2.3) выполняется с $R = (4LM)^{-1}(m^2 - \nu)$, если $\delta \le (2L)^{-1}(m^2 - \nu)$, $0 < \nu < m^2$.

Пусть $\tilde{x}^*$ есть произвольная точка глобального минимума $\tilde{J}$ на D. Тогда $\|F(\tilde{x}^*) - \tilde{f}\|_{H_2} \le \|F(x) - \tilde{f}\|_{H_2}$ для всех $x \in D$. Полагая здесь $x = x^*$ и используя (1.12), получаем

$$\|\tilde{x}^* - x^*\|_{H_1} \le \frac{2\delta}{m}. \tag{2.4}$$

Следствие 2. Пусть выполнены условия теоремы 1 и условие

$$\delta \le \frac{1}{2} Rm.$$

Тогда множество $\tilde{X}^* = \{\tilde{x}^*\}$ решений задачи (1.9) одноточечное. В области сильной выпуклости $D \cap O_R(x^*)$ функционала $\tilde{J}$ лежит единственная его стационарная точка $\tilde{x}^*$ и имеет место оценка (2.4).

2) Пусть $p \in (1,2]$. В этом случае локальная сильная выпуклость функционала невязки уже не может быть гарантирована даже при $\tilde{f} = f$. Тем не менее в случае точных данных для функционала

$$J(x) = \frac{1}{2}\|F(x) - f\|^2_{H_2} \tag{2.5}$$

имеет место

Теорема 2. *Пусть* $h = \delta = 0$ *и выполняется условие* 1.

1. *Предположим, что* $p \in (1,2)$. *Тогда для всех* $x \in D \cap O_R(x^*)$ *при*

$$R = \frac{1}{M}\left(\frac{2(1-\varepsilon)m^{2/p}}{L}\right)^{p/(2-p)}, \quad \varepsilon \in (0,1), \tag{2.6}$$

справедливо

$$(J'(x), x - x^*)_{H_1} \ge \varepsilon m^2 \|x - x^*\|^{2p}_{H_1}. \tag{2.7}$$

2. *Пусть* $p = 2$ *и выполняется условие*

$$L \le 2(1-\varepsilon)m, \quad \varepsilon \in (0,1). \tag{2.8}$$

Тогда для всех $x \in D$ *с указанным значением* p *имеет место оценка* (2.7).

Доказательство. Для произвольных $x \in D$, $p \in (1,2]$ с использованием (1.4) получаем

$$(J'(x), x - x^*)_{H_1} = (F(x) - f, F'(x)(x - x^*))_{H_2} \ge \|F(x) - f\|^2_{H_2} - \frac{1}{2} L \|F(x) - f\|_{H_2} \|x - x^*\|^2_{H_1}.$$

Отсюда с учетом (1.12) следует

$$(J'(x), x - x^*)_{H_1} \ge \|F(x) - f\|^2_{H_2}\left(1 - \frac{L}{2m^{2/p}}\|F(x) - f\|^{(2-p)/p}_{H_2}\right). \tag{2.9}$$

Используя (1.3), (2.6), (2.8), на основании (2.9) заключаем, что при выполнении условий теоремы справедливо

$$(J'(x), x - x^*)_{H_1} \ge \varepsilon \|F(x) - f\|^2_{H_2}. \tag{2.10}$$

Требуемая оценка непосредственно следует из (1.12), (2.10). Теорема доказана.

Точка $x^* = F^{-1}(f)$, доставляющая глобальный минимум функционалу J на D, является стационарной для J. Следующее утверждение устанавливает отсутствие других стационарных точек J в окрестности x^*.

Следствие 3. В условиях п. 1 теоремы 2 функционал J не имеет в $D \cap O_R(x^*)$ других стационарных точек кроме x^*. В условиях п. 2 теоремы 2 x^* является единственной стационарной точкой J на всем множестве D.

Для доказательства заметим, что если $\hat{x}$ — стационарная точка функционала (2.5), то согласно (1.10) $(J'(\hat{x}), \hat{x} - x^*)_{H_1} \le 0$. Требуемое утверждение непосредственно следует теперь из (2.7).

В следующем примере покажем, что результат следствия 3 не допускает непосредственного обобщения на случай $p > 2$ и, кроме того, условие (2.8) не может быть ослаблено.

Пример 1. Рассмотрим инъективное отображение $F : D \to \mathbb{R}$, $D = [-a, b]$, $a, b > 0$, $F(x) = b^p - (b - x)^p$, $p \geq 2$. Непосредственно устанавливается оценка

$$|F(x) - F(y)| \geq |x - y|^p, \quad x, y \in D, \tag{2.11}$$

неулучшаемая по порядку $|x - y|$. Кроме того,

$$|F'(x) - F'(y)| \leq p(p-1)(a+b)^{p-2}|x - y|, \quad x, y \in D. \tag{2.12}$$

Пусть $f = 0$, тогда $x^* = 0$. Функция $J(x) = 0.5F^2(x)$ имеет на D стационарные точки $\hat{x}^{(1)} = 0$, $\hat{x}^{(2)} = b$. Как видно из (2.11), (2.12), в этом примере оценки (1.2), (1.3) и (1.12) выполняются с $L = p(p-1)(a+b)^{p-2}$, $M = p(a+b)^{p-1}$ и $m = 1$ соответственно. Пусть вначале $p > 2$. Выбирая $R = b$ произвольно малым и варьируя $a > 0$, видим, что окрестность $O_R(x^*) \cap D$ содержит две стационарные точки $\hat{x}^{(1)} = 0$, $\hat{x}^{(2)} = b$ функции J на D. Таким образом, следствия 1—3 непосредственно не обобщаются на случай, когда условие 1 выполнено с показателем $p > 2$. Полагая $p = 2$, имеем $L = 2$, так что условие (2.8) не выполнено ни при каком $\varepsilon > 0$. Вновь для любого $R = b > 0$ имеем две стационарные точки в $O_R(x^*) \cap D$. Следовательно, условие (2.8) в п. 2 следствия 3 не допускает ослабления.

3. МЕТОД ПРОЕКЦИИ ГРАДИЕНТА И КЛАСТЕРИЗАЦИЯ СТАЦИОНАРНЫХ ТОЧЕК ФУНКЦИОНАЛА НЕВЯЗКИ

Далее нас будет интересовать поведение метода проекции градиента

$$x_{n+1} = P_D(x_n - \rho\tilde{J}'(x_n)), \quad x_0 \in D, \quad \rho > 0, \tag{3.1}$$

в применении к задаче (1.9) в предположении, что $h + \delta > 0$. Здесь P_D — оператор метрического проектирования из H_1 на D. Анализу итераций (3.1) предпошлем несколько примеров, иллюстрирующих вид $\tilde{J}$ и множества стационарных точек $\tilde{J}$ на D в случае приближенного оператора. Следующий пример показывает, что если $p = 1$, то при $h > 0$ в отличие от случая $h = 0$ (теорема 1) сильная выпуклость $\tilde{J}$ на D в наших предположениях может отсутствовать.

Пример 2. Положим $F(x) = x$, $\tilde{F}(x) = x + hn^{-1}\sin(nx)$, $\tilde{f} = f = 0$, $D = [0,1]$. Условия (1.2), (1.12) очевидно выполнены с $L = 0$, $m = 1$, $p = 1$. Условия (1.6), (1.7) также выполнены. Имеем

$$\tilde{J}''(x) = (1 + h\cos(nx))^2 - hn(x + hn^{-1}\sin(nx))\sin(nx).$$

Отсюда следует, что при $x = x_m = (2n)^{-1}\pi + 2\pi n^{-1}m$, $m \in \mathbb{Z}$, будет $\tilde{J}''(x_m) = 1 - h(\pi/2 + h) - 2\pi mh$. Выбирая $h = n^{-1/2}$, $m = [n^{3/4}]$, $n \in \mathbb{N}$, получаем $0 < x_m \leq (2n)^{-1}\pi + 2\pi n^{-1/4}$. Поэтому точки x_m лежат в D и сколь угодно близки к $x^* = 0$ при достаточно больших n. Кроме того, для указанных n с подходящей константой $\kappa_0 > 0$ имеем $\tilde{J}''(x_m) \leq -\kappa_0 n^{1/4} < 0$. Таким образом, сильная выпуклость $\tilde{J}$ нарушается в сколь угодно малой окрестности решения x^*. Пример 2 закончен.

Эффект разрушения сильной выпуклости функционала невязки при $p = 1$ проиллюстрируем также ниже в примере 4.

Если $p \in (1,2]$, то при возмущениях вида (1.6), (1.7) множество стационарных точек $\tilde{J}$ на D может иметь достаточно сложный вид.

Пример 3. В координатном гильбертовом пространстве

$$H_1 = \left\{ \sum_{i=1}^{\infty} x_i e_i : \sum_{i=1}^{\infty} x_i^2 < \infty \right\},$$

определенном ортонормированным базисом $\{e_i\}_{i=1}^{\infty}$, введем оператор $F: H_1 \to H_2 = H_1 \times \mathbb{R}^N$:

$$F(x) = (Ax; G(x)), \quad G(x) = (x_1^2, \ldots, x_N^2),$$

где $A^* = A \ge O$ – самосопряженный вполне непрерывный оператор с нулевым ядром и полным набором собственных векторов $\{e_i\}_{i=1}^{\infty}$. Обозначим через $\{\lambda_i\}_{i=1}^{\infty}$ соответствующие положительные собственные значения, $\lim\limits_{i\to\infty} \lambda_i = 0$. Норма в H_2 определяется равенством $\|(x;y)\|_{H_2}^2 = \|x\|_{H_1}^2 + \|y\|_{\mathbb{R}^N}^2$, $(x;y) \in H_2$. Положим также $f = (0;0)$,

$$D = \{x = A^q v, \ \|v\|_{H_1} \le d\}, \quad q \ge 1, \quad d > 0.$$

Ясно, что D компактно, оператор F инъективен на D и решением задачи (1.1) является элемент $x^* = 0$. В координатном представлении оператор A имеет вид

$$Ax = \sum_{i=1}^{\infty} \lambda_i x_i e_i, \quad x = \sum_{i=1}^{\infty} x_i e_i.$$

Будем считать, что λ_i малы при $1 \le i \le N$, $i \ge M+1$, где $M > N$, и

$$\max_{1\le i\le N} \lambda_i \le \max_{i \ge M+1} \lambda_i.$$

Определим приближенный оператор

$$\tilde{F}(x) = (\tilde{A}x; G(x)), \quad \tilde{A}x = \sum_{i=N+1}^{M} \lambda_i x_i e_i.$$

В данном случае

$$\max_{x\in D} \|\tilde{F}(x) - F(x)\|_{H_2} \le \|\tilde{A} - A\|_{L(H_1,H_1)} d(\max_{i\ge 1} \lambda_i)^q,$$
$$\|\tilde{F}'(x) - F'(x)\|_{L(H_1,H_2)} = \|\tilde{A} - A\|_{L(H_1,H_1)}.$$

Кроме того, $\|\tilde{A} - A\|_{L(H_1,H_1)} = \max\limits_{i\ge M+1} \lambda_i$. Таким образом, условия (1.6), (1.7) выполняются с $h = \sigma \max\limits_{i\ge M+1} \lambda_i$, $\sigma = \max\{1, d(\max\limits_{i\ge 1} \lambda_i)^q\}$. Положим $\tilde{f} = (0; \delta/\sqrt{N}, \ldots, \delta/\sqrt{N})$, так что $\|\tilde{f} - f\|_{H_2} = \delta$. Непосредственно проверяется, что условие Липшица (1.2) выполнено с $L = 2$. Для проверки условия 1 запишем

$$\|F(x) - F(y)\|_{H_2} = (\|A(x-y)\|_{H_1}^2 + \|G(x) - G(y)\|_{\mathbb{R}^N}^2)^{1/2} \ge \|A(x-y)\|_{H_1}. \tag{3.2}$$

Поскольку $x, y \in D$, имеем $x = A^q u$, $y = A^q v$, где $\|u\|_{H_1}, \|v\|_{H_1} \le d$. Таким образом,

$$x - y = A^q w, \quad \|w\|_{H_1} = \|u - v\|_{H_1} \le 2d.$$

Используя неравенство моментов (см. [11, гл. 2, § 8]), далее получаем

$$\|x-y\|_{H_1} = \|A^q w\|_{H_1} \le \|A^{q+1} w\|_{H_1}^{q/(q+1)} \|w\|_{H_1}^{1/(q+1)} \le \|A(x-y)\|_{H_1}^{q/(q+1)} (2d)^{1/(q+1)}.$$

Отсюда с учетом (3.2) следует

$$\|F(x) - F(y)\|_{H_2} \ge (2d)^{-1/q} \|x - y\|_{H_1}^{(q+1)/q}, \quad x, y \in D.$$

Видим, что условие 1 выполняется с $m = (2d)^{-1/q}$ и показателем $p = (q+1)/q$. При соответствующем выборе $q \ge 1$ величина p принимает любые значения в $(1,2]$. Функционал $\tilde{J}$ в рассматриваемом случае имеет вид

$$\tilde{J}(x) = \frac{1}{2} \sum_{i=N+1}^{M} \lambda_i^2 x_i^2 + \frac{1}{2} \sum_{i=1}^{N} (x_i^2 - \delta/\sqrt{N})^2.$$

Поскольку

$$(\tilde{J}'(x),\eta)_{H_1} = \sum_{i=N+1}^{M} \lambda_i^2 x_i \eta_i + 2\sum_{i=1}^{N}(x_i^2 - \delta/\sqrt{N})x_i\eta_i, \quad \eta \in H_1,$$

имеем следующий набор множеств в H_1, на которых $\tilde{J}'(x) = 0$:

$$\mathcal{M}(\bar{x}_1,\ldots,\bar{x}_N) = \left\{(\bar{x}_1,\ldots,\bar{x}_N,0,\ldots,0,x_{M+1},x_{M+2},\ldots) : \sum_{i=M+1}^{\infty} x_i^2 < \infty\right\}.$$

Здесь $\bar{x}_i \in \{0,\pm\sqrt{\delta}/\sqrt[4]{N}\}, 1 \le i \le N$, выбираются произвольно. Таким образом, градиент $\tilde{J}'$ обращается в нуль на 3^N непересекающихся множествах $\mathcal{M}(\bar{x}_1,\ldots,\bar{x}_N)$. Нетрудно указать условия, при которых все эти множества имеют общие точки с D. Достаточное условие для $\mathcal{M}(\bar{x}_1,\ldots,\bar{x}_N) \cap D \ne \varnothing$ для всех 3^N наборов $\{\bar{x}_1,\ldots,\bar{x}_N\}$ имеет вид

$$\frac{\delta}{\sqrt{N}}\sum_{i=1}^{N}\frac{1}{\lambda_i^{2q}} \le \Delta^2 < d^2. \tag{3.3}$$

Здесь множество стационарных точек $\tilde{J}$ на D содержит семейство $\hat{X}^*$ из 3^N дизъюнктных компактов коразмерности M, полученных параллельным переносом компакта

$$\mathcal{K} = \left\{x \in H_1 : x_i = 0, 1 \le i \le M, \sum_{i=M+1}^{\infty}\frac{x_i^2}{\lambda_i^{2q}} \le d^2 - \Delta^2\right\} \tag{3.4}$$

в точки $(\bar{x}_1,\ldots,\bar{x}_N,0,0,\ldots)$. Оценим расстояние найденных фрагментов множества стационарных точек $\tilde{J}$ на D до решения x^*. Используя (3.3), (3.4), получаем, что для любой точки $x \in \hat{X}^*$ выполняется

$$\|x\|_{H_1}^2 = \sum_{i=1}^{N}x_i^2 + \sum_{i=M+1}^{\infty}x_i^2 \le \sqrt{N}\delta + \sum_{i=M+1}^{\infty}x_i^2 \le \sigma^{-2q}\Delta^2 h^{2q} + \sigma^{-2q}(d^2-\Delta^2)h^{2q} = \sigma^{-2q}d^2h^{2q}.$$

Таким образом, для всех $h \in (0,h_0]$ с фиксированным h_0 имеет место оценка

$$\operatorname{dist}_{H_1}(x^*,\hat{X}^*) \le dh^q \le C_1 h^{1/p}. \tag{3.5}$$

Пример 3 закончен.

Через $C_1, C_2, \ldots$ обозначаем фиксированные постоянные, не зависящие от h, μ, δ. В (3.5) и далее используется обозначение $\operatorname{dist}_{H_1}(x,\Omega) = \inf\{\|x-y\|_{H_1} : y \in \Omega\}$, $\Omega \subset H_1$.

Следующий пример иллюстрирует появление обширного кластера стационарных точек в окрестности решения в случае $p = 1$ в условиях (1.6), (1.8).

Пример 4. В пространстве H_1 из примера 3 рассмотрим компакт

$$D = \left\{x = \sum_{i=1}^{\infty}x_i e_i : |x_i| \le \frac{1}{2^i}, i = 1,2,\ldots\right\}$$

и тождественное отображение $F(x) = x$, $x \in H_1 = H_2$. В данном случае (1.2) и условие 1 выполнены очевидным образом, при этом $m = 1$, $p = 1$, $L = 0$. Положим $f = 0$, тогда $x^* = 0$. В качестве приближенных данных выберем

$$\tilde{F}(x) = \sum_{i=1}^{N}x_i e_i + \sum_{i=N+1}^{\infty}x_i^2 e_i, \quad \tilde{f} = \sum_{i=N+1}^{\infty}\delta_i e_i, \quad 0 < \delta_i \le \frac{1}{4^i}, \quad i \ge N+1.$$

Тогда

$$\|\tilde{f} - f\|_{H_1} \le \delta = \frac{1}{4^N\sqrt{15}}.$$

Кроме того, при $x \in D$ справедливо

$$\|\tilde{F}(x) - F(x)\|_{H_1}^2 = \sum_{i=N+1}^{\infty} (x_i^2 - x_i)^2 \le \sum_{i=N+1}^{\infty} \frac{1}{2^{2(i-1)}} = \frac{1}{3\times 4^{N-1}}.$$

Таким образом, условие (1.6) выполняется при $h = 1/(2^{N-1}\sqrt{3})$. Оценим величину μ в условии (1.8). Имеем

$$(\tilde{F}'(x) - F'(x))(x - x^*) = \sum_{i=N+1}^{\infty} x_i(2x_i - 1)e_i,$$

$$\|(\tilde{F}'(x) - F'(x))(x - x^*)\|_{H_1}^2 = \sum_{i=N+1}^{\infty} x_i^2(2x_i - 1)^2 \le 4\sum_{i=N+1}^{\infty} x_i^2 \le \frac{1}{3\times 4^{N-1}}.$$

Поэтому можно принять $\mu = h$. В рассматриваемом случае

$$\tilde{J}(x) = \frac{1}{2}\sum_{i=1}^{N} x_i^2 + \frac{1}{2}\sum_{i=N+1}^{\infty} (x_i^2 - \delta_i)^2, \quad (\tilde{J}'(x), \eta)_{H_1} = \sum_{i=1}^{N} x_i\eta_i + 2\sum_{i=N+1}^{\infty} (x_i^2 - \delta_i)x_i\eta_i, \quad \eta \in H_1. \tag{3.6}$$

Из (3.6) следует, что $\tilde{J}'(x) = 0$ на множестве точек

$$\hat{X}^* = \left\{x = \sum_{i=N+1}^{\infty} \bar{x}_i e_i : \bar{x}_i \in \{0, \pm\sqrt{\delta_i}\},\ i \ge N+1\right\},$$

имеющем мощность континуума. Указанные точки являются стационарными для $\tilde{J}$ на D. Непосредственно проверяется, что

$$\operatorname{dist}_{H_1}(x^*, \hat{X}^*) \le \frac{1}{2^N\sqrt{3}} = \frac{1}{2}h. \tag{3.7}$$

Пример 4 закончен.

Отсутствие свойства (сильной) выпуклости минимизируемого функционала значительно затрудняет анализ итераций (3.1) и других итерационных методов решения задачи (1.9) (см. [10]). Известные утверждения о скорости сходимости этих методов по норме получены в предположении о сильной выпуклости целевого функционала. Теоремы 1, 2 утверждают единственность стационарной точки функционала (1.9) в окрестности решения x^* при $h = 0$ или $h = \delta = 0$ в шаре, диаметр которого определяется параметрами исходной задачи (1.1). Примеры 3, 4 показывают, что при замене точных данных (F, f) их приближенными аналогами $(\tilde{F}, \tilde{f})$ указанное свойство локальной одноэкстремальности в общем случае не имеет места. В целом распределение стационарных точек функционала невязки в случае приближенных данных приобретает следующий вид. Вместо единственной точки глобального минимума функционала J будем иметь кластер стационарных точек $\tilde{J}$ в окрестности x^*. Существенно, что все точки кластера находятся в шаре $O_{r(h,\mu,\delta)}(x^*)$ радиуса

$$r(h, \mu, \delta) = O((h + \mu + \delta)^{1/p}). \tag{3.8}$$

Отмеченное свойство кластеризации в случае $h = \mu = 0$ было ранее обосновано в [12], [13]. Для доказательства утверждения в общем виде ниже покажем, что множество $B_{r(h,\mu,\delta)} = O_{r(h,\mu,\delta)}(x^*) \cap D$ является аттрактором итераций (3.1) в том смысле, что

$$\lim_{n\to\infty} \operatorname{dist}_{H_1}(x_n, B_{r(h,\mu,\delta)}) = 0. \tag{3.9}$$

Одновременно будет получена оценка скорости сходимости в (3.9), заменяющая классические оценки скорости сильной сходимости градиентных итераций в сильно выпуклых задачах. Основное условие выполнения равенства (3.9) имеет вид $x_0 \in O_{r_0}(x^*) \cap D$, где величина $r_0 = r_0(L, M_0, m)$ определяется параметрами задачи (1.1). Поскольку последовательность (3.1) стационарна по n при старте из любой стационарной точки функционала $\tilde{J}$, из (3.9) следует, что все стационарные точки задачи (1.9), принадлежащие $O_{r_0}(x^*) \cap D$, лежат исключительно в аттракторе $B_{r(h,\mu,\delta)}$ диаметра $O((h + \mu + \delta)^{1/p})$. Совокупность описанных результатов определяет принцип кластеризации стационарных точек задачи (1.9) в случае возмущений общего вида. Заметим, что оценки (3.5), (3.7) из примеров 3, 4 согласуются по порядку с размером кластера, указанном в (3.8).

Перейдем к обоснованию анонсированных выше утверждений об асимптотических свойствах итераций (3.1).

4. ОЦЕНКА АТТРАКТОРА ИТЕРАЦИЙ МЕТОДА ПРОЕКЦИИ ГРАДИЕНТА

В этом разделе будем считать выполненными условие 1 и соотношения (1.5), (1.6), (1.8). Исследуем сходимость итераций (3.1). Используя свойство

$$\|P_D(x) - P_D(y)\|_{H_1} \le \|x - y\|_{H_1} \quad \forall x, y \in H_1,$$

из (3.1) получаем

$$\|x_{n+1} - x^*\|_{H_1}^2 \le \|x_n - x^*\|_{H_1}^2 - 2\rho(\tilde{J}'(x_n), x_n - x^*)_{H_1} + \rho^2 \|\tilde{J}'(x_n)\|_{H_1}^2. \tag{4.1}$$

С учетом (1.11) запишем

$$\begin{gathered}(\tilde{J}'(x_n), x_n - x^*)_{H_1} = (F(x_n) - f, F'(x_n)(x_n - x^*))_{H_2} + (\tilde{F}(x_n) - F(x_n) + f - \tilde{f}, F'(x_n)(x_n - x^*))_{H_2} + \\ + (\tilde{F}(x_n) - F(x_n) + f - \tilde{f}, (\tilde{F}'(x_n) - F'(x_n))(x_n - x^*))_{H_2} + (F(x_n) - f, (\tilde{F}'(x_n) - F'(x_n))(x_n - x^*))_{H_2}.\end{gathered} \tag{4.2}$$

Для первого скалярного произведения в правой части (4.2) в силу (1.4) имеем оценку

$$\begin{gathered}(F(x_n) - f, F'(x_n)(x_n - x^*))_{H_2} = (F(x_n) - f, F(x_n) - f + G(x_n, x^*))_{H_2} \ge \\ \ge \|F(x_n) - f\|_{H_2}^2 - \frac{1}{2} L \|F(x_n) - f\|_{H_2} \|x_n - x^*\|_{H_1}^2.\end{gathered} \tag{4.3}$$

Сумма остальных скалярных произведений с учетом (1.4), (1.5), (1.8) оценивается сверху величиной

$$(h + \mu + \delta)\|F(x_n) - f\|_{H_2} + \frac{1}{2} L(h + \delta)\|x_n - x^*\|_{H_1}^2 + \mu(h + \delta). \tag{4.4}$$

Из условия 1 следует, что

$$\|x_n - x^*\|_{H_1}^2 \le C_2 \|F(x_n) - f\|_{H_2}, \quad C_2 = (d_0^{p-2} m)^{-1}.$$

Поэтому сумма первого и второго слагаемых в (4.4) мажорируется величиной

$$C_3(h + \mu + \delta)\|F(x_n) - f\|_{H_2} \le \frac{1}{2}\|F(x_n) - f\|_{H_2}^2 + C_4(h + \mu + \delta)^2. \tag{4.5}$$

Здесь $C_3 = 1 + C_2 L/2$, $C_4 = C_3^2/2$. Из (4.2)–(4.5) следует

$$(\tilde{J}'(x_n), x_n - x^*)_{H_1} \ge \frac{1}{2}\|F(x_n) - f\|_{H_2}^2 - \frac{1}{2} L \|F(x_n) - f\|_{H_2} \|x_n - x^*\|_{H_1}^2 - (C_4 + 1)(h + \mu + \delta)^2. \tag{4.6}$$

С использованием (1.5), (1.6), (1.8), (1.11) находим

$$\|\tilde{J}'(x_n)\|_{H_1} \le \|\tilde{F}'(x_n)\|_{L(H_1, H_2)} \|\tilde{F}(x_n) - \tilde{f}\|_{H_2} \le M_0(\|F(x_n) - f\|_{H_2} + h + \delta),$$

поэтому

$$\|\tilde{J}'(x_n)\|_{H_1}^2 \le 2M_0^2 \|F(x_n) - f\|_{H_2}^2 + 2M_0^2(h + \delta)^2. \tag{4.7}$$

Объединяя неравенства (4.1), (4.6), (4.7), получаем оценку

$$\begin{gathered}\|x_{n+1} - x^*\|_{H_1}^2 \le \|x_n - x^*\|_{H_1}^2 - \rho\|F(x_n) - f\|_{H_2}^2 + L\rho\|F(x_n) - f\|_{H_2} \|x_n - x^*\|_{H_1}^2 + \\ + 2M_0^2\rho^2 \|F(x_n) - f\|_{H_2}^2 + 2(C_4 + 1)\rho(h + \mu + \delta)^2 + 2M_0^2\rho^2(h + \delta)^2 \le \\ \le \|x_n - x^*\|_{H_1}^2 - \rho\|F(x_n) - f\|_{H_2}^2 + L\rho\|F(x_n) - f\|_{H_2} \|x_n - x^*\|_{H_1}^2 + \\ + 2M_0^2\rho^2 \|F(x_n) - f\|_{H_2}^2 + K(\rho)\rho(h + \mu + \delta)^2, \quad K(\rho) = 2(C_4 + 1) + 2M_0^2\rho.\end{gathered}$$

Отсюда с учетом (1.12) получаем

$$\|x_{n+1} - x^*\|_{H_1}^2 \le \|x_n - x^*\|_{H_1}^2 - \rho\|F(x_n) - f\|_{H_2}^2 + \frac{L\rho}{m^{2/p}}\|F(x_n) - f\|_{H_2}^{(2+p)/p} + \\ + 2M_0^2\rho^2\|F(x_n) - f\|_{H_2}^2 + K(\rho)\rho(h + \mu + \delta)^2. \tag{4.8}$$

Перепишем (4.8) в виде

$$\|x_{n+1} - x^*\|_{H_1}^2 \le \|x_n - x^*\|_{H_1}^2 - \\ - \rho\|F(x_n) - f\|_{H_2}^2\left(1 - \frac{L}{m^{2/p}}\|F(x_n) - f\|_{H_2}^{(2-p)/p} - 2M_0^2\rho\right) + K(\rho)\rho(h + \mu + \delta)^2. \tag{4.9}$$

Сделаем индуктивное предположение о том, что

$$\frac{L}{m^{2/p}}\|F(x_n) - f\|_{H_2}^{(2-p)/p} + 2M_0^2\rho \le \frac{1}{2}. \tag{4.10}$$

Тогда (4.9) вместе с (1.12) влечет оценку

$$\|x_{n+1} - x^*\|_{H_1}^2 \le \|x_n - x^*\|_{H_1}^2 - \frac{1}{2}\rho m^2\|x_n - x^*\|_{H_1}^{2p} + C_5\rho(h + \mu + \delta)^2, \quad C_5 = K(1/(4M_0^2)). \tag{4.11}$$

Для удобства далее будем считать выполненными следующие условия, достаточные для выполнения (4.10):

$$\|x_n - x^*\|_{H_1} \le \omega \triangleq \frac{1}{M_0}\left(\frac{m^{2/p}}{4L}\right)^{p/(2-p)}, \quad p \in [1,2); \quad \frac{L}{m} \le \frac{1}{4}, \quad p = 2; \tag{4.12}$$

$$0 < \rho \le \frac{1}{8M_0^2}. \tag{4.13}$$

Покажем теперь, что если $p \in [1,2)$, то неравенство $\|x_n - x^*\|_{H_1} \le \omega$ из (4.12) при подходящих дополнительных условиях выполняется также для $x = x_{n+1}$. Рассмотрим два возможных случая.

1) Пусть $p \in (1,2)$. Нетрудно убедиться, что функция

$$\varphi(\lambda) = \lambda - \frac{1}{2}\rho m^2\lambda^p$$

монотонно возрастает при $\lambda \in [0,\tilde{\lambda}]$, где

$$\tilde{\lambda} = \left(\frac{2}{pm^2\rho}\right)^{1/(p-1)}.$$

Предположим, что

$$\rho < \frac{2}{pm^2\omega^{2p-2}}. \tag{4.14}$$

Тогда $\|x_n - x^*\|_{H_1}^2 \le \omega^2 \le \tilde{\lambda}$, поэтому

$$\varphi(\|x_n - x^*\|_{H_1}^2) \le \varphi(\omega^2). \tag{4.15}$$

2) Пусть $p = 1$. Тогда при выполнении (4.14) функция $\varphi(\lambda)$ монотонно возрастает при $\lambda \ge 0$. Следовательно, неравенство (4.15) также имеет место.

В дополнение предположим, что

$$h + \mu + \delta \le \frac{m\omega^p}{\sqrt{2C_5}}. \tag{4.16}$$

При выполнении (4.14), (4.16) из (4.11) следует

$$\|x_{n+1} - x^*\|_{H_1}^2 \le \varphi(\omega^2) + C_5\rho(h + \mu + \delta)^2 = \omega^2 - \frac{1}{2}\rho m^2\omega^{2p} + C_5\rho(h + \mu + \delta)^2 \le \omega^2.$$

Таким образом, при выполнении условия

$$\|x_0 - x^*\|_{H_1} \le \omega \tag{4.17}$$

будем иметь $\|x_n - x^*\|_{H_1} \le \omega$ для всех $n = 1,2,\ldots$. Поэтому оценка (4.11) справедлива для всех $n = 0,1,\ldots$. Следующая теорема подытоживает проведенные рассуждения.

Теорема 3. *Пусть выполняются условие* 1 *и условия* (1.5), (1.6), (1.8), (4.13). *Предположим, что в случае* $p = 2$ *выполняется* $L/m \le 1/4$, *а в случае* $p \in [1,2)$ *имеют место* (4.14), (4.16), (4.17). *Тогда для* $a_n = \|x_n - x^*\|^2_{H_1}$ *выполняется оценка*

$$a_{n+1} \le a_n - \frac{1}{2}\rho m^2 a_n^p + C_5\rho(h + \mu + \delta)^2, \quad n = 0,1,\ldots. \tag{4.18}$$

Покажем теперь, что множество $B_{r(h,\mu,\delta)} = O_{r(h,\mu,\delta)}(x^*) \cap D$ с $r(h,\mu,\delta)$ вида (3.8) является аттрактором итераций (3.1) и дадим оценку скорости сходимости в соотношении (3.9). В случае $p = 1$ из (4.18) следует

$$a_{n+1} \le \left(1 - \frac{1}{2}\rho m^2\right)a_n + C_5\rho(h + \mu + \delta)^2. \tag{4.19}$$

Применяя к оценке (4.19) лемму 1 из [14, гл. 2, § 2], получаем следующее утверждение.

Теорема 4. *Пусть* $p = 1$, *выполняются условие* 1 *и условия* (1.5), (1.6), (1.8), (4.13), (4.14), (4.16), (4.17). *Тогда*

$$\|x_n - x^*\|^2_{H_1} \le \frac{2C_5}{m^2}(h + \mu + \delta)^2 + \|x_0 - x^*\|^2_{H_1}\left(1 - \frac{1}{2}\rho m^2\right)^n, \quad n = 0,1,\ldots. \tag{4.20}$$

Следствие 4. В условиях теоремы 4 справедливо

$$\operatorname{dist}_{H_1}(x_n, B_{r(h,\mu,\delta)}) \le \|x_0 - x^*\|_{H_1}\left(1 - \frac{1}{2}\rho m^2\right)^{n/2}, \quad r(h,\mu,\delta) = \frac{\sqrt{2C_5}}{m}(h + \mu + \delta).$$

Пусть теперь $p \in (1,2]$. Могут представиться три случая.

1) Для всех $n = 0,1,\ldots$ справедливо

$$a_n \ge C_6(h + \mu + \delta)^{2/p}, \quad C_6 = \left(\frac{2C_5}{m^2}\right)^{1/p}. \tag{4.21}$$

В этом случае $C_5\rho(h + \mu + \delta)^2 \le \rho m^2 a_n^p/2$, поэтому (4.18) влечет $0 \le a_{n+1} \le a_n$, $n = 0,1,\ldots$. Следовательно, существует предел $\overline{a} = \lim\limits_{n\to\infty} a_n$. Переходя в (4.18) к пределу, получаем $\overline{a} \le C_6(h + \mu + \delta)^{2/p}$. Поэтому из (4.21) следует

$$\lim_{n\to\infty} a_n = \overline{a} = C_6(h + \mu + \delta)^{2/p}. \tag{4.22}$$

2) Существует такой номер n_0, что

$$a_n \ge C_6(h + \mu + \delta)^{2/p}, \quad n = 0,\ldots,n_0; \quad a_n \le C_6(h + \mu + \delta)^{2/p}, \quad n \ge n_0 + 1.$$

Как и в случае 1), получаем $a_{n+1} \le a_n$ для всех $n = 0,1,\ldots,n_0$. Вместо (4.22) здесь имеем $\limsup\limits_{n\to\infty} a_n \le C_6(h + \mu + \delta)^{2/p}$.

3) Для некоторого номера n_1 выполняется

$$a_{n_1} \le C_6(h + \mu + \delta)^{2/p}, \quad a_{n_1+1} > C_6(h + \mu + \delta)^{2/p}.$$

Тогда из (4.18) следует

$$a_{n_1+1} \le a_{n_1} + C_5\rho(h + \mu + \delta)^2 \le C_7(h + \mu + \delta)^{2/p}, \quad C_7 = C_6 + C_5\rho\Delta_0^{(2p-2)/p}. \tag{4.23}$$

Как и в случае 1), $a_{n_1+2} \le a_{n_1+1}$, поэтому $a_{n_1+2} \le C_7(h + \mu + \delta)^{2/p}$. Далее, если $a_{n_1+2} > C_6(h + \mu + \delta)^{2/p}$, то, как и выше, $a_{n_1+3} \le a_{n_1+2} \le C_7(h + \mu + \delta)^{2/p}$. Если же $a_{n_1+2} \le C_6(h + \mu + \delta)^{2/p}$, то, аналогично (4.23),

$a_{n_1+3} \le C_7(h+\mu+\delta)^{2/p}$. Поэтому в случае 3) для всех номеров $n \ge n_1$ выполняется $a_n \le$ $\le C_7(h+\mu+\delta)^{2/p}$.

Проведенный анализ показывает, что во всех случаях 1)–3) справедливо

$$\limsup_{n\to\infty} a_n \le C_7(h+\mu+\delta)^{2/p}.$$

Более того, последовательность a_n является монотонно убывающей до момента достижения ею величины $C_7(h+\mu+\delta)^{2/p}$, а после этого момента $a_n \le C_7(h+\mu+\delta)^{2/p}$. Таким образом, при $r(h,\mu,\delta) = C_8(h+\mu+\delta)^{1/p}$, $C_8 = \sqrt{C_7}$ имеет место (3.9).

Для получения оценки скорости сходимости в (3.9) зафиксируем произвольно $\kappa \in (0,1/2)$ и заметим, что условие

$$a_n \ge C_9(h+\mu+\delta)^{2/p}, \quad C_9 = \left(\frac{2C_5}{(1-2\kappa)m^2}\right)^{1/p} \tag{4.24}$$

влечет

$$-\frac{1}{2}\rho m^2 a_n^p + C_5\rho(h+\mu+\delta)^2 \le -\kappa\rho m^2 a_n^p.$$

Поэтому ввиду (4.18) имеем

$$a_{n+1} \le a_n - \kappa\rho m^2 a_n^p. \tag{4.25}$$

Следовательно, если (4.24) выполняется для номеров $n = 0,1,\ldots,l$, то для этих же номеров n справедлива оценка (4.25). Используя [14, гл. 2, § 2, лемма 6], получаем

$$a_n \le \frac{a_0}{\left(1+(p-1)a_0^{p-1}\kappa\rho m^2 n\right)^{1/(p-1)}}, \quad n = 0,1,\ldots,l. \tag{4.26}$$

Применим оценку (4.26) к номерам n таким, что $a_n > C_{10}(h+\mu+\delta)^{2/p}$, $C_{10} = \max\{C_7, C_9\}$. Получаем следующую оценку скорости сходимости итераций (3.1) к аттрактору $B_{r(h,\mu,\delta)}$, $r(h,\mu,\delta) = C_{11}(h+\mu+\delta)^{1/p}$, $C_{11} = \sqrt{C_{10}}$.

Теорема 5. *Пусть выполняются условие* 1 *и условия* (1.5), (1.6), (1.8). *Предположим, что в случае* $p \in (1,2]$ *выполняются* (4.13), (4.14), (4.16), (4.17), *а при* $p = 2$ *имеет место* $L/m \le 1/4$. *Тогда*

$$\limsup_{n\to\infty} \|x_n - x^*\|_{H_1} \le C_8(h+\mu+\delta)^{1/p}.$$

Кроме того, справедлива оценка

$$\|x_n - x^*\|_{H_1} \le \max\left\{\frac{\|x_0 - x^*\|_{H_1}}{\left(1+(p-1)\|x_0 - x^*\|_{H_1}^{2p-2}\kappa\rho m^2 n\right)^{1/(2p-2)}}, C_{11}(h+\mu+\delta)^{1/p}\right\}. \tag{4.27}$$

Следствие 5. В условиях теоремы 5 справедливо

$$\operatorname{dist}_{H_1}(x_n, B_{r(h,\mu,\delta)}) \le \frac{\|x_0 - x^*\|_{H_1}}{\left(1+(p-1)\|x_0 - x^*\|_{H_1}^{2p-2}\kappa\rho m^2 n\right)^{1/(2p-2)}},$$
$$r(h,\mu,\delta) = C_{11}(h+\mu+\delta)^{1/p}.$$

Следствия 4 и 5 приводят к следующему результату о локализации кластера стационарных точек функционала (1.9) на множестве D.

Теорема 6. *Пусть выполняются условие* 1 *и условия* (1.5), (1.6), (1.8), (4.16). *Тогда все стационарные точки функционала* $\tilde{J}$ *на* D, *принадлежащие* $O_\omega(x^*) \cap D$, *лежат в* $B_{r(h,\mu,\delta)}$, $r(h,\mu,\delta) = O((h+\mu+\delta)^{1/p})$.

Теорема 6 уточняет более грубую оценку размера кластера стационарных точек в условиях (1.5)–(1.7) (см. [15]), согласно которой кластер лежит в шаре $O_{r(h,\delta)}(x^*)$ радиуса $r(h,\delta) = O(\sqrt{h} + \delta)^{1/p}$.

Теоремы 4, 5 не содержат информации о сходимости последовательности $\{x_n\}$ к какому-либо пределу. Положительные утверждения такого рода получаются при наложении на приближенный оператор $\tilde{F}$ и шаговый множитель ρ подходящих дополнительных условий. Предположим, что производная $\tilde{F}'$, наряду с F', удовлетворяет условию Липшица (1.2). В этом случае градиент $\tilde{J}'$ также удовлетворяет условию Липшица:

$$\|\tilde{J}'(x) - \tilde{J}'(y)\|_{H_1} \le \Lambda\|x - y\|_{H_1} \quad \forall x, y \in D.$$

Из теоремы 1.4.4 (см. [10]) следует, что при выполнении дополнительного условия

$$0 < \rho < \frac{2}{\Lambda} \tag{4.28}$$

для итераций (3.1) справедливо $\lim\limits_{n\to\infty}\|x_{n+1} - x_n\|_{H_1} = 0$. С другой стороны, в силу компактности D любое бесконечное подмножество последовательности $\{x_n\}$ имеет сильно сходящуюся подпоследовательность $\{x_{n_k}\}$, $\lim\limits_{k\to\infty}\|x_{n_k} - \bar{x}\|_{H_1} = 0$. Тогда и $\lim\limits_{k\to\infty}\|x_{n_k+1} - \bar{x}\|_{H_1} = 0$. Полагая в (3.1) $n = n_k$ и переходя к пределу при $k \to \infty$, заключаем, что $\bar{x} = P_D(\bar{x} - \rho\tilde{J}'(\bar{x}))$, т.е. $\bar{x}$ есть стационарная точка $\tilde{J}$ на D. Тем самым получаем следующее уточнение теорем 4 и 5.

Теорема 7. *Пусть выполняются условия теорем* 4, 5*, производная* $\tilde{F}'$ *удовлетворяет условию Липшица* (1.2) *и выполнено* (4.28). *Тогда последовательность* $\{x_n\}$ *имеет предельные точки, и каждая такая точка* $\bar{x}$ *является стационарной для* $\tilde{J}$ *на* D. *При этом* $\bar{x} \in B_{r(h,\mu,\delta)}$ *с* $r(h,\mu,\delta)$*, указанными в следствиях* 4 *и* 5.

5. О НЕОБХОДИМОСТИ УСЛОВИЯ 1

Покажем необходимость условия 1 для теорем 4, 5 и их следствий. Ради упрощения изложения примем $h = \mu = 0$. Согласно следствиям 4 и 5, при выполнении надлежащих условий на характеристики задачи (1.1) и параметры процесса (3.1) имеет место равенство (3.9), т.е.

$$\lim_{n\to\infty} \operatorname{dist}_{H_1}(x_n, B_{r(\delta)}) = 0, \quad r(\delta) = C_{12}\delta^{1/p}. \tag{5.1}$$

Здесь $x_n = x_n(\tilde{f})$ – порождаемая (3.1) последовательность. Одним из условий для (5.1) является $\|x_0 - x^*\|_{H_1} \le \omega$, величина ω определена в (4.12). Это условие выполнено для любого $x_0 \in D$, если $\operatorname{diam}_{H_1}(D) \le \omega$. Оценки скорости сходимости из теорем 4 и 5 позволяют указать момент останова итераций $n = n(\delta)$ так, что оператор $R_\delta(\tilde{f}) = x_{n(\delta)}(\tilde{f})$ определяет регуляризующий алгоритм для задачи (1.1). Это означает, что

$$\lim_{\delta\to 0}\Delta_D(\delta) = 0, \quad \Delta_D(\delta) \triangleq \sup_{x\in D, \|\tilde{f} - F(x)\|_{H_2}\le\delta} \|R_\delta(\tilde{f}) - x\|_{H_1}. \tag{5.2}$$

Величина $\Delta_D(\delta)$ есть равномерная на D оценка точности алгоритма R_δ на входных данных, известных с погрешностью δ. Используя оценки скорости сходимости (4.20), (4.27), нетрудно выбрать момент останова $n = n(\delta)$ так, что в условиях теорем 4, 5 для погрешности $\Delta_D(\delta)$ справедлива оценка

$$\Delta_D(\delta) \le C_{13}\delta^{1/p}. \tag{5.3}$$

Отвлекаясь от оценок скорости сходимости $\{x_n\}$ к аттрактору $B_{r(\delta)}$, на основании (5.1) констатируем, что оператор итерационного перехода

$$T(\tilde{f})(x) = P_D(x - \rho\tilde{J}'(x)), \quad \tilde{J}(x) = \frac{1}{2}\|F(x) - \tilde{f}\|^2_{H_2},$$

обеспечивает асимптотическую регуляризацию задачи (1.1) в смысле предельного соотношения

$$\limsup_{n\to\infty}\left\|T^n(\tilde{f})(x_0)-x\right\|_{H_1}\le r(\delta)\quad \forall x_0\in D,\quad \forall x\in D,\quad \left\|\tilde{f}-F(x)\right\|_{H_2}\le\delta. \tag{5.4}$$

Соотношение (5.4) формализует факт устойчивой сходимости итераций $x_{n+1}=T(\tilde{f})(x_n)$ к аттрактору $B_{r(\delta)}$. Однако, если $T(\tilde{f}):D\to H_1$ – произвольное отображение со свойством (5.4), то при отсутствии равномерной по $\tilde{f}$ оценки скорости сходимости в (5.4) отображение $\tilde{f}\to T^n(\tilde{f})$ ни с каким фиксированным правилом останова $n=n(\delta)$ в общем случае не порождает регуляризующий алгоритм для (1.1) в смысле (5.2). Тем не менее следующая теорема утверждает, что без привлечения условия 1 построение итерационных операторов $T(\tilde{f})$, обладающих достаточно слабым свойством асимптотической регуляризации (5.4), невозможно, каково бы ни было $p\ge 1$.

Теорема 8. *Пусть D – связное компактное множество,*

$$\delta_0=\operatorname{diam}_{H_2}(F(D)),$$

$F:D\to H_2$ – инъективный оператор. Пусть для некоторого оператора $T(\tilde{f}):D\to H_1$ при всех $\delta\in(0,\delta_0]$ выполняется (5.4) *с $r(\delta)=C\delta^{1/p}$, $p\ge 1$. Тогда*

$$\left\|F(x)-F(y)\right\|_{H_2}\ge m_0\left\|x-y\right\|_{H_1}^p\quad \forall x,y\in D;\quad m_0=(2C)^{-p}. \tag{5.5}$$

Доказательство. Зафиксируем $x_1^*,x_2^*\in D$ так, что для $f_j=F(x_j^*)$, $j=1,2$, выполняется $\left\|f_1-f_2\right\|_{H_2}=\delta\in(0,\delta_0]$. Поскольку $F(D)$ компактно и связно и $\delta\le\operatorname{diam}_{H_2}(F(D))$, элементы $f_j\in F(D)$, $j=1,2$, со свойством $\left\|f_1-f_2\right\|_{H_2}=\delta$ существуют. Поэтому $x_j^*=F^{-1}(f_j)$ также определены для всех $\delta\in(0,\delta_0]$. Рассмотрим уравнение (1.1) с $f=f_1$, решением которого является $x^*=x_1^*$. В качестве приближенной правой части $\tilde{f}$ выберем $\tilde{f}=f_1$. В данном случае $\left\|\tilde{f}-f\right\|_{H_2}=0\le\delta$. Рассмотрим также задачу (1.1) с $f=f_2$ и приближенным элементом $\tilde{f}=f_1$. Ее точным решением является $x^*=x_2^*$. В этом случае $\left\|\tilde{f}-f\right\|_{H_2}=\delta\le\delta$. Согласно (5.4), для произвольного $x_0\in D$ и соответствующих положительных последовательностей $\{\alpha_n\}$, $\{\beta_n\}$ выполняется

$$\begin{gathered}\|T^n(f_1)(x_0)-x_1^*\|_{H_1}\le C\delta^{1/p}+\alpha_n,\quad \lim_{n\to\infty}\alpha_n=0,\\ \|T^n(f_1)(x_0)-x_2^*\|_{H_1}\le C\delta^{1/p}+\beta_n,\quad \lim_{n\to\infty}\beta_n=0.\end{gathered} \tag{5.6}$$

Из (5.6) следует

$$\|x_1^*-x_2^*\|_{H_1}\le 2C\delta^{1/p}+\alpha_n+\beta_n,\quad n=0,1,\ldots.$$

Переходя здесь к пределу при $n\to\infty$, получаем

$$\|x_1^*-x_2^*\|_{H_1}\le 2C\delta^{1/p}=2C\|F(x_1^*)-F(x_2^*)\|_{H_2}^{1/p}. \tag{5.7}$$

Поскольку любые пары точек (x_1^*,x_2^*) соответствуют некоторым $f_j=F(x_j^*)$, $j=1,2$, со свойством $\left\|f_1-f_2\right\|_{H_2}\le\delta$, $\delta\in(0,\delta_0]$, неравенство (5.7) имеет место для любых $x^*,x_2^*\in D$. Соотношение (5.5) непосредственно следует из (5.7). Теорема доказана.

Замечание 1. Повторяя доказательство теоремы 8 с $\alpha_n=\beta_n=0$ и заменой $T^n(f_1)(x_0)$ на $R_\delta(f_1)$, убеждаемся, что условие 1 является также необходимым для существования регуляризующего алгоритма с оценкой (5.3). В случае $p=1$ этот факт был ранее отмечен в [16].

Полагая $p=1$ и объединяя теоремы 1 и 8, с учетом замечания 1 получаем следующий вариант необходимого критерия регуляризуемости задачи (1.1) с оценкой точности $O(\delta)$.

Следствие 6. Пусть D — связное компактное множество, $F\colon D \to H_2$ — инъективный оператор, $\delta_0 = \mathrm{diam}_{H_2}(F(D)) \le ((2C)^{-2p} - \nu)/(2L)$, $\nu > 0$. Предположим, что для некоторого регуляризующего алгоритма $R_\delta(\tilde{f})$ выполняется оценка $\Delta_D(\delta) \le C\delta$, $\delta \in (0, \delta_0]$. Тогда для любого $f \in F(D)$ функционал невязки $\tilde{J}$ является сильно выпуклым в $D \cap O_R(x^*)$, $R = (4LM)^{-1}((2C)^{-2p} - \nu)$, $x^* = F^{-1}(f)$.

Теорема 8 показывает, что свойство асимптотической регуляризуемости задачи (1.1) в смысле (5.4) и формально более сильное свойство ее регуляризуемости в смысле классического определения (5.2) с оценкой (5.3) по существу эквивалентны при любом $p \ge 1$.

СПИСОК ЛИТЕРАТУРЫ

1. *Треногин В.А.* Функциональный анализ. М.: Наука, 1980.
2. *Бакушинский А.Б., Кокурин М.Ю.* Алгоритмический анализ нерегулярных операторных уравнений. М.: ЛЕНАНД, 2012.
3. *Лаврентьев М.М., Романов В.Г., Шишатский С.П.* Некорректные задачи математической физики и анализа. М.: Наука, 1980.
4. *Кабанихин С.И.* Обратные и некорректные задачи. Новосибирск: Сиб. науч. изд-во, 2008.
5. *Романов В.Г.* Устойчивость в обратных задачах. М.: Научный мир, 2004.
6. *Isakov V.* Inverse Problems for Partial Differential Equations. N.Y.: Springer, 2006.
7. *Кокурин М.Ю.* Об условно корректных и обобщенно корректных задачах // Ж. вычисл. матем. и матем. физ. 2013. Т. 53. № 6. С. 857–866.
8. *Kokurin M.Yu.* On a characteristic property of conditionally well-posed problems // J. Inv. Ill-Posed Probl. 2015. V. 23. № 3. P. 245–262.
9. *Тихонов А.Н., Леонов А.С., Ягола А.Г.* Нелинейные некорректные задачи. М.: Наука, Физматлит, 1995.
10. *Васильев Ф.П.* Методы решения экстремальных задач. М.: Наука, 1981.
11. *Красносельский М.А., Вайникко Г.М., Забрейко П.П., Рутицкий Я.Б., Стеценко В.Я.* Приближенное решение операторных уравнений. М.: Наука, 1969.
12. *Kokurin M.Yu.* On stable finite dimensional approximation of conditionally well-posed inverse problems // Inv. Probl. 2016. V. 32. № 10. 105007.
13. *Kokurin M.Yu.* Stable gradient projection method for nonlinear conditionally well-posed inverse problems // J. Inv. Ill-posed Probl. 2016. V. 24. № 3. P. 323–332.
14. *Поляк Б.Т.* Введение в оптимизацию. М.: Наука, 1983.
15. *Кокурин М.Ю.* О кластеризации стационарных точек функционалов невязки условно-корректных обратных задач // Сиб. журн. вычисл. матем. 2018. Т. 21. № 4. С. 393–406.
16. *Леонов А.С.* О возможности получения линейных оценок точности приближенных решений обратных задач // Изв. вуз. Матем. 2016. № 10. С. 29–35.