Малинов В.Г.

Ульяновский государственный университет

О РЕГУЛЯРИЗОВАННОМ ПРОЕКЦИОННОМ ЭКСТРАГРАДИЕНТНОМ ДВУХТОЧЕЧНОМ МЕТОДЕ МИНИМИЗАЦИИ С ПЕРЕМЕННОЙ МЕТРИКОЙ

(Редактирована 07.11.2019 г.)

A regularization method in Hilbert space for problems of minimization with inaccurate initial date, based on the iterative projection extragradient two-point first order variable metric method in conjunction with the Tikhonov function method is proposed. Sufficient conditions for the convergence and the rate of convergence of the method are proved.

Исследуется метод регуляризации для решения задач минимизации с неточными исходными данными на простом выпуклом замкнутом множестве в сепарабельном гильбертовом пространстве, основанный на проекционном обобщённом двухточечном двухэтапном экстраградиентном методе (ПОДЭМ) с переменной метрикой. Обосновываются достаточные условия сходимости, оценка скорости сходимости, правило останова; для базового метода приводятся результаты сравнительных численных экспериментов с методами первого и второго порядка.

1. Постановка задачи.

Множество математических моделей науки и техники приводят к задаче минимизации на выпуклом замкнутом множестве Q

$$f(\mathbf{x}) \to inf, \quad \mathbf{x} \in Q \subset H, \tag{1}$$

где сепарабельное гильбертово пространство H нормировано скалярным произведением, $\|\mathbf{x}\| = (\mathbf{x}, \mathbf{x})^{1/2}$ $\forall$ $\mathbf{x} \in H$; функция $f(\mathbf{x})$ определена, выпукла и непрерывно дифференцируема по Фреше на Q, её градиенты (обозначаем часто штрихом) Липшицевы:

$$\exists \quad L = const > 0, \quad \|\nabla f(\mathbf{x}) - \nabla f(\mathbf{u})\| \le L\|\mathbf{x} - \mathbf{u}\|, \quad \mathbf{x}, \mathbf{u} \in Q \ . \tag{2}$$

Предполагаем, что: поверхности уровней функции $f(\mathbf{x})$ "овражные";

$$inf\, f(\mathbf{x}) = f_* > -\infty,\ \mathbf{x} \in Q; \quad Q_* = \{\mathbf{x} \in Q:\ f(\mathbf{x}) = f_*\} \neq \varnothing; \qquad (3)$$

градиенты целевой функции $f(\mathbf{x})$ имеют возмущения. Первое предположение говорит о необходимости применения для решения задачи специальных методов, приспособленных для минимизации "овражных" функций;

а последнее — о неустойчивости задачи и необходимости для ее решения методов регуляризации [1]–[5].

2. **Подходы к решению задачи и построение метода**

Для решения поставленной задачи в последние десятилетия предложены и исследованы итеративные и непрерывные проекционные методы минимизации (НПММ) (см., например, [6]–[23]). Здесь лишь кратко отметим, что НПММ позволяют использовать развитый аппарат вычислительной математики для решения дифференциальных уравнений [12], но представляют практический интерес и их итеративные аналоги. Итеративные методы первого порядка (в том числе многошаговые), разработанные для решения описанной задачи, обладают характерным свойством, почти все замедляются в окрестности минимума, чем и объясняется их линейная скорость сходимости. От этого недостатка свободны методы переменной метрики (МПМ). Поэтому, на основе подхода из работ [15] к построению методов минимизации, предложены и исследованы *непрерывные методы* проекции градиента (НМПГПМ) с переменной метрикой первого и второго порядков (см. [15]–[17]).

На основе подхода к построению непрерывных НМПГПМ, для решения задачи минимизации на выпуклом замкнутом множестве Q в евклидовом пространстве E^n, в работах [18]–[21], [23] построены и исследованы НПММ переменной метрики (НПММПМ) с их итеративными аналогами и регуляризованными версиями. Из итеративных это проекционные обобщённые двухточечные экстраградиентные МПМ, кратко (ПОДЭМПМ) назовём их квазиньютоновскими (ПОДЭМК) в случае аппроксимации гессиана или обращённого гессиана . Предложен и исследован ПОДЭМК с постоянными параметрами метода, в [12] исследован ПОДЭМ для задач минимизации с ограничениями равенствами и неравенствами. Здесь исследуется регуляризованный ПОДЭМК для решения задачи (1) с неточными исходными данными *на простом выпуклом замкнутом множестве* в сепарабельном гильбертовом пространстве H.

Для построения метода, наряду с имеющимися скалярным произведением и проекцией $P_Q(\mathbf{v})$ в точке $\mathbf{x} \in H$ вектора $\mathbf{v} \in H$ на множество Q в этой метрике, в сепарабельном гильбертовом пространстве H введём новую метрику с помощью нового скалярного произведения $(\mathbf{B}(\mathbf{x})\mathbf{u}, \mathbf{u})$, аналогично тому, как это сделано в [15], [16]. $\mathbf{B}(\mathbf{x})$: $H \to H$ при каждом фиксированном $\mathbf{x} \in H$ является положительно определённым линейным самосопряженным оператором метрики пространства. Критерием проекции $\mathbf{w} \in Q$ вектора $\mathbf{v} \in H$ на множество Q в новой метрике служит неравенство

$$(\mathbf{B}(\mathbf{x})(\mathbf{w} - \mathbf{v}), \mathbf{u} - \mathbf{w}) \geq 0, \quad \mathbf{u} \in Q. \tag{4}$$

В построенном пространстве для решения задачи (1) исследуем регуляризованный ПОДЭМК (градиент часто обозначаем штрихом):

$$\text{(1 этап)} \quad \mathbf{y}^k = \mathbf{x}^k - \mathbf{x}^{k-1}; \ \mathbf{z}^k = \mathbf{x}^k + \alpha_k \ \mathbf{y}^k;$$

$$\text{(2 этап)} \quad \mathbf{x}^{k+1} = P_Q[\mathbf{z}^k - \beta_k \mathbf{B}(\mathbf{z}^k)^{-1} \nabla \mathbf{t}_k(\mathbf{z}^k)], \ k = 0,1,2,\ldots, \tag{5}$$

где $\mathbf{x}^0$ – произвольная начальная точка из H; $\mathbf{x}^{-1} = \mathbf{x}^0$; $\alpha_k > 0, \quad \beta_k > 0$ – параметры метода; $\mathbf{B}(\mathbf{z}^k)^{-1}$ – оператор, обратный к $\mathbf{B}(\mathbf{x})$ в точке $\mathbf{z}^k$;

$$\nabla \mathbf{t}_k(\mathbf{z}^k) = \nabla \mathbf{f}_k(\mathbf{z}^k) + \tau_k \mathbf{z}^k, \quad k \geq 0 \ - \tag{6}$$

приближение в точке $\mathbf{z}^k$ точного градиента $\nabla T(\mathbf{x}) = \nabla f(\mathbf{x}) + \tau \mathbf{x}$ функции Тихонова $T(\mathbf{x}) = f(\mathbf{x}) + \tau \|\mathbf{x}\|^2 / 2$; приближённые градиенты таковы, что

$$\max_{\mathbf{x} \in Q} \|\nabla f_k(\mathbf{x}) - \nabla f(\mathbf{x})\| \leq \delta_k (1 + \|\mathbf{x}\|) \ \forall \ k \geq 0. \tag{7}$$

В силу предположений, функция $T(\mathbf{x})$ дифференцируема по Фреше на множестве Q; полагаем, что выполнено свойство её непрерывной дифференцируемости на Q.

3. **Вычислительные способности базового метода**

Кратко остановимся на свойствах базового метода минимизации, на котором построен регуляризованный метод (5), (6). Им служит предложенный и исследованный в [11] для решения задачи (1) итеративный ПОДЭМК с параметрами α_k и γ_k

$$\text{(1 этап)} \quad \mathbf{y}^k = \mathbf{x}^k - \mathbf{x}^{k-1}, \;\; \mathbf{z}^k = \mathbf{x}^k + \alpha_k \mathbf{y}^k;$$

$$\text{(2 этап)} \quad \mathbf{x}^{k+1} = P_Q[\mathbf{z}^k - \gamma_k \mathbf{B}(\mathbf{z}^k)^{-1} \nabla f(\mathbf{z}^k)], \quad k = 0,1,2,\dots,$$

где $\forall\ \mathbf{x}^0 \in E^n$ – начальная точка; $\mathbf{x}^{-1} = \mathbf{x}^0$; $\mathbf{B}(\mathbf{z}^k) = \mathbf{B}_k$ – последовательность положительно определённых самосопряженных линейных операторов, задающих метрику в каждой точке $\mathbf{z}^k \in E^n$; $\mathbf{B}(\mathbf{z}^k)^{-1} = \mathbf{B}_k^{-1}$ – обратные к $\mathbf{B}_k$ операторы. Заметим, что итерационной формулой задается семейство методов переменной метрики, так как каждый способ выбора параметров метода и оператора $\mathbf{B}(\mathbf{x})$ определяет новый метод. Исследован его частный случай, ПОДЭМПМ второго порядка (TSVM) с формулой

$$\mathbf{x}^{k+1} = P_Q[\mathbf{z}^k - \beta_k \nabla^2 \mathbf{f}(\mathbf{z}^k)^{-1} \nabla f(\mathbf{z}^k)], \quad k = 0,1,2,\dots,$$

соответствующей выбору оператора $\mathbf{B}(\mathbf{x}) = \mathbf{f}''(\mathbf{x})\ \forall\ \mathbf{x} \in Q$; в этих методах $P_Q(\mathbf{v})$ – проекция точки $\mathbf{v}$ на множество Q; $\alpha_k, \beta_k > 0$ – параметры метода; $\nabla^2 \mathbf{f}(\mathbf{z}^k)^{-1}$ – матрица, обратная матрице Гессе функции $f(\mathbf{x})$ $\forall$ $\mathbf{z}^k \in E^n$; $\mathbf{x}^{-1} = \mathbf{x}^0 \in E^n$ [21].

TSVM по теоретической оценке скорости сходимости (квадратичная) и точности решения задачи (1) превзошел все другие методы, включая лучшие одношаговые методы переменной метрики первого порядка Давидона-Флетчера-Пауэлла (DFP) и Бройдена-Флетчера-Гольдфарба-Шенно (BFGS). Обобщённые двухшаговые МПМ первого порядка также численно исследуются, их программы отлаживаются.

Вычислительные эксперименты *иллюстрируют возможности базового метода* для исследуемого, имеющего меньшую скорость сходимости. Они проведены на *известных тестовых задачах минимизации* различных размерностей с "овражными" линиями, поверхностями или гиперповерхностями уровней различных конфигураций, в том числе:

1) Б.Т. Поляка и В.А. Скокова на основе овражной функции Розенброка,

$$f(\mathbf{x}) = 100(\mathbf{x}_2 - \mathbf{x}_1^2)^2 + (1 - \mathbf{x}_1)^2, \quad \mathbf{x} \in [-\mathbf{3};\, \mathbf{3}] \in E^2;$$

2) на основе функции с поверхностями уровней, имеющими форму геликоидных оврагов:

$$f(\mathbf{x}) = 100\left[(x_1 - 10\cos x_3)^2 + 10(x_2 - \sin x_3)^2\right] + 100(\pi - \mathbf{x}_3)^2, \mathbf{x} \in [-\mathbf{11};\, \mathbf{5}] \in E^3;$$

3) на основе функции

$$f(\mathbf{x}) = 100\left[x_n - \left(\sum_{i=1}^{i=n-1} x_i^2\right)^2 / (n-1)\right]^2 + \sum_{i=1}^{i=n-1} (1 - x_i)^2,$$

с "овражными" поверхностями уровней, $\mathbf{x} \in [-\mathbf{4};\, \mathbf{4}] \in E^n$, ($n \geq 2$, чётные);

4) на основе овражной функции

$$f(\mathbf{x}) = 100\sum_{i=2}^{i=n} (\mathbf{x}_i - \mathbf{x}_1^2)^2 + (1 - \mathbf{x}_1)^2, \quad \mathbf{x} \in [-\mathbf{4};\, \mathbf{4}] \in E^n \, (n = 2, 4, 6, 8,....).$$

В таблице приведены результаты решения этих задач для TSVM в сравнении с другими, ставшими классическими, методами минимизации первого и второго порядка приведены в таблице. В ней присутствуют следующие сокращённые обозначения методов первого и второго порядка:

Desk – метод наискорейшего спуска;

New1 – первая модификация модифицированного метода Ньютона (ММН); New2 – вторая модификация ММН; New3 – третья модификация ММН;

DFP, BFGS – квазиньютоновские (см. выше);

MFR – сопряжённых градиентов Флетчера-Ривса;

PPR – метод сопряжённых градиентов Б.Т. Поляка - Полака-Рибьера;

MS – сопряжённых градиентов Соренсена;

GAM – метод П.Р. Гамбурда; LED – метод Le D [25] в нашей проекционной версии; GFPM – обобщённый четырёхпараметрический проекционный двухточечный [10];

GTPM – обобщённый двухпараметрический двухточечный, получаемый из TSVM при $\nabla^2 f(\mathbf{z}^k) = \mathbf{I}$.

Другие обозначения в таблице**:** qit – количество итераций;

qg – количество вычислений градиента;

qcf – количество вычислений значения функции;

f_{qit} – полученное методом значение функции, при останове;

$\delta x = max\ |x_i^{qit} - x_i^*|,\ i \in \{1{:}n\}$; значок ***** в столбце значений функции означает, что метод не решил задачу.

Останов методов происходил по одному из признаков:

1) норма градиента меньше заданного числа, $nrg < \varepsilon$ = 1.e-5;

2) значения функции совпали на трёх последовательных итерациях с точностью ε_1 = 1.0e-14;

3) превышено заданное число итераций метода. Матрица Гессе вычислялась численно, через градиенты.

Таблица результатов численных экспериментов.

	№ задачи – n	Метод	qit	qcf	qg	f_{qit}	δx
1	**1) -** 2	Desk	100	763	100	*0.2492	0.750
	$f(\mathbf{x}^0) = 24.2$,	New1	13	145	65	1.9765d-17	4.d-8
	$\mathbf{x}^0 = (-1.2;\ 1.0)$,	New2	12	98	61	3.3105d-12	2.d-6
	$\mathbf{x}^* = (1.0;\ 1.0)$,	New3	14	181	70	0	0
	$f(\mathbf{x}^*) = 0.$	TSVM	11	127	28	1.6730d-26	3.7d-14
		DFP	32	159	157	5.5621d-24	4.1d-12
		BFGS	34	161	159	1.5650d-26	2.1d-12
		GFPM	24	345	26	3.9404e-15	1.2e-7
		GTPM	14	383	23	7.5494e-12	5.3e-6
		MS	35	293	35	1.4211e-12	4.0e-6
		GAM	26	505	28	1.0668e-8	2.0e-4
		LED	13	336	15	2.3306e-12	2.0e-6
2	2) – 3	Desk	5	54	6	*7.56511	7.7660
	$f(\mathbf{x}^0) = 1000009.87$,	New1	49	686	344	5.5769d-8	0.0045
	$\mathbf{x}^0 = (0;\ 0;\ 0)$,	New2	50	404	350	1.4139d-10	0.0010
	$\mathbf{x}^* = (-10, 0; \pi)$,	New3	51	728	357	1.2746d-9	0.0016
	$f(\mathbf{x}^*) = 0.$	TSVM	39	404	272	4.4294d-14	3.02d-9
		DFP	25	148	146	*80.80	8.481

		BFGS	12	101	99	*9.047	20.03
3	3) – 48	Desk	100	574	100	*1.7546	0.269
	$f(\mathbf{x}^0) = 121.208$,	New1	6	59	583	5.5478d-15	1.4d-8
	$\mathbf{x}^0 =$(-1.2; 1; ...),	New2	6	66	583	4.1437d-15	.14d-7
	$\mathbf{x}^* = \mathbf{1},\ f(\mathbf{x}^*) = 0.$	New3	6	56	583	2.0133d-13	1.1d-8
		TSVM	5	98	105	1.6764d-16	4.6d-8
		TSVM	6	116	299	3.0815d-29	2.d-14
		DFP	28	134	132	3.9147d-19	1.2d-10
		BFGS	30	127	125	8.0160d-18	4.4d-9
4	4) – 48	Desk	100	750	100	*0.5002	0.914
	$f(\mathbf{x}^0) = 16499.56$,	New1	39	383	3784	2.5532d-16	.32d-7
	$\mathbf{x}^0 =$(-1.2; 1; ...),	New2	37	317	3589	5.729d-16	1.1d-9
	$\mathbf{x}^* = \mathbf{1},\ f(\mathbf{x}^*) = 0.$	New3	36	356	3492	*1.3138d-4	1.d-14
		TSVM	37	299	1086	2.5265d-20	.32d-9
		TSVM	29	352	1378	4.3245d-18	3.8d-9
		DFP	30	159	157	*1.1810	1.06
		BFGS	25	134	132	*0.8460	.905

Отметим, что в работе [15] и других с помощью численных экспериментов исследованы многошаговые МПМ первого порядка *безусловной минимизации*, обоснований сходимости и оценок скорости сходимости не проведено. Полученные в исследованиях оценки скорости сходимости [11] и результаты численных экспериментов характеризуют ПОДЭМК как очень перспективные и ценные для приложений.

4. **Обоснование сходимости метода (5), (6)**

Нормальным решением задачи (1) называют её решение с минимальной нормой**;** исследуем достаточные условия сходимости к нормальному решению задачи последовательности $\left\{\mathbf{x}^k\right\}$ регуляризованного метода (5), (6).

Теорема 1. *Пусть выполнены следующие условия:*

1) множество $Q \subset H$ выпукло и замкнуто, H сепарабельно;

2) функция $f(\mathbf{x}) \in C^{1,1}(Q)$ выпуклая, справедливы соотношения (3);

3) для приближений $\nabla f_k(\mathbf{x})$ точного градиента $\nabla f(\mathbf{x})$ выполнено (7);

4) оператор $\mathbf{B}(\mathbf{x})\ \forall\ \mathbf{x} \in H$ таков, что

$$m\|\mathbf{u}\|^2 \le (\mathbf{B}(\mathbf{x})\mathbf{u},\mathbf{u}) \le M\|\mathbf{u}\|^2, \quad 0 < m \le M,\ \mathbf{x},\mathbf{u} \in H; \tag{8}$$

5) существует выпуклая функция $\varphi(\mathbf{x}) \in C^{1,1}(Q)$ *такая, что её градиент*

$$\nabla\varphi(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x})\nabla f(\mathbf{x}) \quad \forall\ \mathbf{x} \in H;$$

6) параметры $\alpha_k, \beta_k, \tau_k, \delta_k$ *метода* (5), (6) *таковы, что:*

$$\alpha_k \ge \alpha_{k+1} > 0,\ \beta_k \ge \beta_{k+1} > 0,\ \delta_k \ge 0,\ \tau_k \ge \tau_{k+1} > 0,\ k \ge 0;$$

$$(\tau_k - \tau_{k+1})^2 \left[\beta_{k-1}\tau_{k-1}\right]^{-1} \tau_k^{-2} \to 0, \quad k \to \infty, \tag{9}$$

$$\tau_k - \tau_{k+1} \to 0, \quad \delta_k + \tau_k + \beta_k^2 \left(\delta_k + \tau_k\right)\left(\beta_{k-1}\tau_{k-1}\right)^{-1} \to 0, \quad k \to \infty.$$

Тогда последовательность $\{\mathbf{x}^k\}$, *определяемая методом* (5), (6), (9), *из произвольной начальной точки* $\mathbf{x}^0 \in H$, *равномерно относительно выбора приближённых градиентов* $\nabla f_k(\mathbf{x})$ *в* (7), *по норме* H *сходится к точке* $\mathbf{x}^*$,

$$\left\|\mathbf{x}^k - \mathbf{x}^*\right\| \to 0, \quad k \to \infty, \tag{10}$$

где $\left\|\mathbf{x}^*\right\| = \inf\limits_{\mathbf{x} \in Q_*} \|\mathbf{x}\|$, $\mathbf{x}^* \in Q_*$ – *нормальное решение задачи* (1).

Доказательство. Схема доказательства аналогична использованной при обосновании сходимости других регуляризованных многошаговых методов минимизации (например, из [6]–[8], [10]). Отличие в использовании пятого условия теоремы и эквивалентности метрик, определяемых различными скалярными произведениями в сепарабельном гильбертовом пространстве H. В построенном пространстве, в силу предположений теоремы 1, условия (3) выполнены, множество минимумов Q_* выпукло и замкнуто, нормальное решение задачи (1) и минимум функции $f(\mathbf{x})$ существуют; на $Q \subset H$ существует точка $\mathbf{v}^k$ такая, что (см., например, [6]-[8])

$$\mathrm{T}(\mathbf{v}^k) = \inf_{\mathbf{x} \in Q} T(\mathbf{x}),\ k \ge 0, \quad \lim_{k\to\infty} \left\|\mathbf{v}^k - \mathbf{x}^*\right\| = 0, \tag{11}$$

$$(\nabla T(\mathbf{v}^k), \mathbf{u} - \mathbf{v}^k) \ge 0, \quad \mathbf{u} \in Q, \tag{12}$$

$$\left\|\mathbf{v}^{k+1}-\mathbf{v}^{k}\right\| \le C\tau_k^{-1}\left|\tau_k-\tau_{k+1}\right|, \quad k\ge 0, \tag{13}$$

$$C=\sup_{k\ge 0} max\left\{\left\|\mathbf{v}^{k}\right\|, \left\|\nabla f(\mathbf{v}^{k})\right\|\right\}. \tag{13'}$$

Отметим, что соотношения вида (11)–(13') верны, с разными точками $\mathbf{v}^k$, в пространствах с обоими метриками. Обосновывая (10), оценим правую часть неравенства:

$$\left\|\mathbf{x}^{k}-\mathbf{x}^{*}\right\| \le \left\|\mathbf{x}^{k}-\mathbf{x}^{k+1}\right\| + \left\|\mathbf{x}^{k+1}-\mathbf{v}^{k}\right\| + \left\|\mathbf{v}^{k}-\mathbf{x}^{*}\right\|, \; k\ge 0. \tag{14}$$

Учитывая (11), для правой части (14) покажем выполнение соотношений

$$\left\|\mathbf{x}^{k+1}-\mathbf{x}^{k}\right\| \to 0, \quad \left\|\mathbf{x}^{k+1}-\mathbf{v}^{k}\right\| \to 0, \quad k\to\infty. \tag{15}$$

Из характеристического свойства оператора проектирования ([3], с. 72) и из (5), пользуясь идеей из работы [13], получаем вариационное неравенство

$$(\mathbf{x}^{k+1}-\mathbf{z}^{k}+\beta\,\mathbf{B}^{-1}(\mathbf{z}^{k})\nabla\mathbf{t}_k(\mathbf{z}^{k}), \mathbf{u}-\mathbf{x}^{k+1})\ge 0 \;\; \forall \quad \mathbf{u}\in Q, \quad k\ge 0. \tag{16}$$

(Здесь и далее индекс k у параметров метода часто для краткости опускаем.) Неравенство (12), в силу свойства (4) оператора $\mathbf{B}(\mathbf{x})$ новой метрики в H, эквивалентно равенству

$$\mathbf{v}^{k}=P_Q^{\mathbf{B}(\mathbf{v}^k)}\left[\mathbf{v}^{k}-\beta\mathbf{B}^{-1}(\mathbf{v}^{k})\nabla T(\mathbf{v}^{k})\right], \; \beta>0, \; k\ge 0. \tag{17}$$

Ввиду изоморфности сепарабельных гильбертовых пространств с разными скалярными произведениями ([16], с. 155), в пространстве с исходной метрикой имеет место аналогичное (17), с другой точкой $\mathbf{v}^k$, равенство $\mathbf{v}^{k}=P_Q\left[\mathbf{v}^{k}-\beta\mathbf{B}^{-1}(\mathbf{v}^{k})\nabla T(\mathbf{v}^{k})\right], \; \beta>0, \; k\ge 0$. Отсюда, по свойству оператора проектирования ([3], с.72) в первоначальной метрике пространства H, следует неравенство

$$\beta(\mathbf{B}^{-1}(\mathbf{v}^{k})\nabla T(\mathbf{v}^{k}), \mathbf{u}-\mathbf{v}^{k})\ge 0, \quad \mathbf{u}\in Q, \quad k\ge 0.$$

Положим здесь $\mathbf{u}=\mathbf{x}^{k+1}$, затем сложим с (16), положив в нём $\mathbf{u}=\mathbf{v}^k$:

$$(\mathbf{x}^{k+1}-\mathbf{z}^{k}+\beta\left[\mathbf{B}^{-1}(\mathbf{z}^{k})\nabla\mathbf{t}_k(\mathbf{z}^{k})-\mathbf{B}^{-1}(\mathbf{v}^{k})\nabla T(\mathbf{v}^{k})\right], \mathbf{v}^{k}-\mathbf{x}^{k+1})\ge 0, \tag{18}$$

$\forall \quad k \geq 0$. Подставив выражения для точного и приближённого градиентов функции Тихонова, представим (18) в виде:

$$(\mathbf{x}^{k+1}-\mathbf{z}^k,\mathbf{v}^k-\mathbf{x}^{k+1}) \leq \beta\Big\{\Big(\mathbf{B}^{-1}(\mathbf{z}^k)(\nabla f_k(\mathbf{z}^k)-\nabla f(\mathbf{z}^k)),\mathbf{v}^k-\mathbf{x}^{k+1}\Big)+$$

$$+\Big(\mathbf{B}^{-1}(\mathbf{z}^k)\nabla f(\mathbf{z}^k)-\mathbf{B}^{-1}(\mathbf{v}^k)\nabla f(\mathbf{v}^k)+\tau\mathbf{B}^{-1}(\mathbf{z}^k)(\mathbf{z}^k-\mathbf{v}^k),\mathbf{v}^k-\mathbf{x}^{k+1}\Big)+$$

$$+\tau\Big(\Big[\mathbf{B}^{-1}(\mathbf{z}^k)-\mathbf{B}^{-1}(\mathbf{v}^k)\Big]\mathbf{v}^k,\mathbf{v}^k-\mathbf{x}^{k+1}\Big)\Big\},\ k\geq 0.$$

Преобразуем это неравенство: левую часть – с помощью свойств скалярного произведения, а правую часть – пользуясь третьим, четвертым, пятым условиями теоремы 1, а также (13') и неравенством (см., например, [10])

$$(1/M)\|\mathbf{u}\|^2 \leq (\mathbf{B}^{-1}(\mathbf{x})\mathbf{u},\mathbf{u}) \leq \|\mathbf{u}\|^2/m, \quad \forall \quad \mathbf{x},\mathbf{u}\in Q, \tag{19}$$

где M и m из (8). Тогда придём к неравенству

$$\|\mathbf{x}^{k+1}-\mathbf{v}^k\| + (\mathbf{v}^k-\mathbf{z}^k,\mathbf{x}^{k+1}-\mathbf{v}^k)+\beta\tau\Big(\mathbf{B}^{-1}(\mathbf{z}^k)(\mathbf{z}^k-\mathbf{v}^k),\mathbf{z}^k-\mathbf{v}^k\Big)\leq$$

$$\leq \beta\delta\Big(1+\|\mathbf{z}^k\|\Big)\|\mathbf{x}^{k+1}-\mathbf{v}^k\|/m+\beta\Big(\nabla\varphi(\mathbf{z}^k)-\nabla\varphi(\mathbf{v}^k),\mathbf{v}^k-\mathbf{x}^{k+1}\Big)+ \tag{20}$$

$$+\ \beta\tau\|\mathbf{z}^k-\mathbf{v}^k\|\|\mathbf{v}^k-\mathbf{x}^{k+1}\|/m+2\beta\tau\Big(C\|\mathbf{v}^k-\mathbf{x}^k\|+C\|\mathbf{x}^k-\mathbf{x}^{k+1}\|\Big)/m,\ k\geq 0.$$

Известные неравенства

$$2|ab|\leq \boldsymbol{\varepsilon}a^2+\boldsymbol{\varepsilon}^{-1}b^2,\quad (a+b)^2\leq(1+\varepsilon)a^2+(1+\varepsilon^{-1})b^2,\quad \forall \quad a,b,\varepsilon>0 \tag{21}$$

и (19) позволяют получить оценки слагаемых в (20). Для левой части:

$$(\mathbf{x}^{k+1}-\mathbf{x}^k,\mathbf{v}^k-\mathbf{z}^k)=-\Big(\mathbf{x}^{k+1}-\mathbf{x}^k,\mathbf{x}^k-\mathbf{v}^k\Big)-\alpha\Big(\mathbf{x}^{k+1}-\mathbf{x}^k,\mathbf{y}^k\Big)\geq$$

$$\geq 0.5\Big[(1-\alpha)\|\mathbf{x}^k-\mathbf{x}^{k+1}\|^2+\|\mathbf{x}^k-\mathbf{v}^k\|^2-\|\mathbf{x}^{k+1}-\mathbf{v}^k\|^2-\alpha\|\mathbf{y}^k\|^2\Big];$$

$$(\mathbf{x}^k-\mathbf{v}^k,\mathbf{v}^k-\mathbf{z}^k)=\alpha\Big(\mathbf{v}^k-\mathbf{x}^k,\mathbf{y}^k\Big)-\|\mathbf{x}^k-\mathbf{v}^k\|^2=$$

$$=0.5\Big[\alpha\|\mathbf{v}^k-\mathbf{x}^{k-1}\|^2-(2-\alpha)\|\mathbf{x}^k-\mathbf{v}^k\|^2-\alpha\|\mathbf{y}^k\|^2\Big];$$

$$(\mathbf{x}^{k+1}-\mathbf{v}^k,\mathbf{v}^k-\mathbf{z}^k)=\Big(\mathbf{x}^{k+1}-\mathbf{x}^k,\mathbf{v}^k-\mathbf{z}^k\Big)+\Big(\mathbf{x}^k-\mathbf{v}^k,\mathbf{v}^k-\mathbf{z}^k\Big)\geq$$

$$\ge 0.5\Big[(1-\alpha)\left\|\mathbf{x}^{k+1}-\mathbf{x}^{k}\right\|^{2}+\alpha\left\|\mathbf{x}^{k-1}-\mathbf{v}^{k}\right\|^{2}-\left\|\mathbf{x}^{k+1}-\mathbf{v}^{k}\right\|^{2}- \qquad (22)$$

$$-(1+\alpha)\left\|\mathbf{x}^{k}-\mathbf{v}^{k}\right\|^{2}-2\alpha\left\|\mathbf{y}^{k}\right\|^{2}\Big];$$

$$\left\|\mathbf{z}^{k}-\mathbf{v}^{k}\right\|^{2}\ge(1-\alpha)\left\|\mathbf{x}^{k}-\mathbf{v}^{k}\right\|^{2}-(\alpha-\alpha^{2})\left\|\mathbf{y}^{k}\right\|^{2}.$$

А для правой части (20):

$$\left\|\mathbf{z}^{k}-\mathbf{v}^{k}\right\|^{2}\le(1+\alpha)\left\|\mathbf{x}^{k}-\mathbf{v}^{k}\right\|^{2}+(\alpha+\alpha^{2})\left\|\mathbf{y}^{k}\right\|^{2};$$

$$\left\|\mathbf{x}^{k+1}-\mathbf{v}^{k}\right\|^{2}\le 2\left\|\mathbf{x}^{k+1}-\mathbf{x}^{k}\right\|^{2}+2\left\|\mathbf{x}^{k}-\mathbf{v}^{k}\right\|^{2};$$

$$\left(1+\left\|\mathbf{z}^{k}\right\|\right)\left\|\mathbf{x}^{k+1}-\mathbf{v}^{k}\right\|^{2}\le(C+1)^{2}\beta/2+(\alpha\beta+3\beta+2)\left\|\mathbf{x}^{k}-\mathbf{v}^{k}\right\|^{2}/(2\beta)+$$

$$+(1+\beta)\left\|\mathbf{x}^{k+1}-\mathbf{x}^{k}\right\|^{2}/(2\beta)+(\alpha^{2}+\alpha)\left\|\mathbf{y}^{k}\right\|^{2}/2;$$

$$\left\|\mathbf{z}^{k}-\mathbf{x}^{k+1}\right\|^{2}\le(1+\alpha)\left\|\mathbf{x}^{k}-\mathbf{x}^{k+1}\right\|^{2}+(\alpha+\alpha^{2})\left\|\mathbf{y}^{k}\right\|^{2};$$

$$\left\|\mathbf{z}^{k}-\mathbf{v}^{k}\right\|\left\|\mathbf{z}^{k}-\mathbf{x}^{k+1}\right\|\le 0.5\left(\left\|\mathbf{z}^{k}-\mathbf{v}^{k}\right\|^{2}+\left\|\mathbf{z}^{k}-\mathbf{x}^{k+1}\right\|^{2}\right)\le$$

$$\le 0.5(1+\alpha)\Big[\left\|\mathbf{x}^{k}-\mathbf{x}^{k+1}\right\|^{2}+\left\|\mathbf{x}^{k}-\mathbf{v}^{k}\right\|^{2}+2\alpha\left\|\mathbf{y}^{k}\right\|^{2}\Big];$$

$$C\left\|\mathbf{v}^{k}-\mathbf{x}^{k}\right\|+C\left\|\mathbf{x}^{k}-\mathbf{x}^{k+1}\right\|\le\frac{1}{2\beta}\left(2C^{2}\beta^{2}+\left\|\mathbf{v}^{k}-\mathbf{x}^{k}\right\|^{2}+\left\|\mathbf{x}^{k}-\mathbf{x}^{k+1}\right\|^{2}\right); \qquad (23)$$

$$\left(\nabla\varphi(\mathbf{z}^{k})-\nabla\varphi(\mathbf{v}^{k}),\mathbf{v}^{k}-\mathbf{x}^{k+1}\right)\le L\left\|\mathbf{z}^{k}-\mathbf{x}^{k+1}\right\|^{2}/4.$$

Оценка скалярного произведения (см. [22], с. 175) в последнем неравенстве в (23) справедлива для выпуклой функции $\varphi(\mathbf{x})\in C^{1,1}(Q)$.

Подставив оценки (22), (23) в (20), приходим к неравенству

$$\left\|\mathbf{x}^{k+1}-\mathbf{v}^{k}\right\|^{2}-a_{1}\left\|\mathbf{x}^{k}-\mathbf{v}^{k}\right\|^{2}+a_{2}\left\|\mathbf{x}^{k}-\mathbf{x}^{k+1}\right\|^{2}\le$$

$$\le a_{3}\left\|\mathbf{x}^{k}-\mathbf{v}^{k}\right\|^{2}+a_{4}\left\|\mathbf{y}^{k}\right\|^{2}-\alpha\left\|\mathbf{x}^{k-1}-\mathbf{v}^{k}\right\|^{2}+a_{5},\quad k\ge 0, \qquad (24)$$

где

$$a_{1}=2/5-2(1-\alpha)\beta\tau/M;\quad a_{2}=1-\alpha-\frac{\beta}{2m}[Lm+\alpha(Lm+2\tau)+4\delta+2\tau]-$$

$$-2(\tau+\delta)/m;\quad a_{3}=3/5+\alpha+2(\tau+\delta)/m+\beta[(3+\alpha)\delta+(1+\alpha)\tau]/m;$$

$$a_{4}=2\alpha+2\left(\alpha-\alpha^{2}\right)\beta\tau/M+\left(\alpha^{2}+\alpha\right)\beta(Lm+2\delta+4\tau)/(2m);$$

$$a_5 = C_1\beta^2(\delta+\tau); \quad C_1 = max\left\{(C+1)^2 / m; 4C^2 / m\right\}.$$

Неравенство (24) представим в виде

$$u_k \le (1-s_k)u_{k-1} + w_k + a_5, \quad k-1 \ge N, \qquad (25)$$

где $s_k = \beta_{k-1}\tau_{k-1}$, левая часть u_k равна левой части (24), а w_k равно разности между правой частью (24) без a_5 и первым слагаемым в правой части (25).

Покажем равномерную сходимость последовательности $\left\{u_k\right\}$ относительно выбора приближённых градиентов в (7). Сначала заметим, что существует число $k_0 > 1$ такое, что $\forall$ $k-1 \ge k_0$ с учетом (9) имеют место неравенства:

$$5\beta_k\left\{M[Lm + \alpha_k(Lm + 2\tau_k + 8m\tau_k / M) + 2\tau_k + 4\delta_k] - 8m\tau_k\right\} \le$$

$$\le 2M\left[m - 5m\alpha_k - 10(\tau_k + \delta_k)\right], \quad 5m\alpha_k < m - 10\tau_k - 10\delta_k, \qquad (25')$$

$$\tau_k < (m - 10\delta_k)/10, \quad \delta_k < m/10.$$

Теперь оценим u_k снизу. Из (25) и второго неравенства (21) при $\varepsilon = 1$, при выполнении условий (9) и (25') имеем

$$u_k \ge (1-2a_1)\left\|\mathbf{x}^k - \mathbf{v}^k\right\|^2 + (a_2 - 2a_1)\left\|\mathbf{x}^k - \mathbf{x}^{k+1}\right\|^2 \ge 0, \qquad (26)$$

так как $1-2a_1 \ge 0, \quad a_2 - 2a_1 \ge 0.$

Теперь оценим w_k сверху. В выражении

$$w_k = a_3\left\|\mathbf{x}^k - \mathbf{v}^k\right\|^2 + a_4\left\|\mathbf{y}^k\right\|^2 -$$

$$-\alpha\left\|\mathbf{x}^{k-1} - \mathbf{v}^k\right\|^2 - (1-s_k)\left[\left\|\mathbf{x}^k - \mathbf{v}^{k-1}\right\|^2 - a_{11}\left\|\mathbf{x}^{k-1} - \mathbf{v}^{k-1}\right\|^2 + a_{21}\left\|\mathbf{y}^k\right\|^2\right],$$

где $a_{11} = a_{1(k-1)}$, $a_{21} = a_{2(k-1)}$, воспользуемся следующими из (21) неравен ствами

$$\left\|\mathbf{x}^k - \mathbf{v}^k\right\|^2 \le (1+s_k)\left\|\mathbf{x}^k - \mathbf{v}^{k-1}\right\|^2 + (1+s_k^{-1})\left\|\mathbf{v}^{k-1} - \mathbf{v}^k\right\|^2,$$

$$\left\|\mathbf{x}^{k-1}-\mathbf{v}^{k-1}\right\|^2 \le (1+s_k)\left\|\mathbf{x}^{k-1}-\mathbf{v}^{k}\right\|^2+(1+s_k^{-1})\left\|\mathbf{v}^{k-1}-\mathbf{v}^{k}\right\|^2.$$

Тогда

$$w_k \le \left[a_3+a_{11}(1-s_k)\right](1+s_k^{-1})\left\|\mathbf{v}^{k-1}-\mathbf{v}^{k}\right\|^2+a_6\left\|\mathbf{x}^{k-1}-\mathbf{v}^{k}\right\|^2-$$

$$-\left[1-s_k-a_3(1+s_k)\right]\left\|\mathbf{x}^{k}-\mathbf{v}^{k-1}\right\|^2-\left[a_{21}(1-s_k)-a_4\right]\left\|\mathbf{y}^{k}\right\|^2,$$

где $a_6=a_{11}(1-s_k^2)-\alpha$. Отсюда, с учетом неравенства

$$\left(\mathbf{x}^{k-1}-\mathbf{v}^{k}\right)^2 \le (1+s_k)\left\|\mathbf{x}^{k-1}-\mathbf{v}^{k-1}\right\|^2+(1+s_k^{-1})\left\|\mathbf{v}^{k-1}-\mathbf{v}^{k}\right\|^2,$$

имеем

$$w_k \le a_7\left\|\mathbf{v}^{k-1}-\mathbf{v}^{k}\right\|^2-\left[1-s_k-a_3(1+s_k)\right]\left\|\mathbf{x}^{k}-\mathbf{v}^{k-1}\right\|^2-$$

$$-\left[a_{21}(1-s_k)-a_4\right]\left\|\mathbf{y}^{k}\right\|^2+a_6(1+s_k)\left\|\mathbf{x}^{k-1}-\mathbf{v}^{k-1}\right\|^2,\quad k\ge 0, \tag{27}$$

где $a_7=\left[a_3+a_{11}(1-s_k)+a_6\right](1+s_k^{-1})$. Для последнего слагаемого в (27) из второго неравенства в (21) следует,

$$\left\|\mathbf{x}^{k-1}-\mathbf{v}^{k-1}\right\|^2 \le (1+\varepsilon)\left\|\mathbf{x}^{k-1}-\mathbf{x}^{k}\right\|^2+(1+\varepsilon^{-1})\left\|\mathbf{x}^{k}-\mathbf{v}^{k-1}\right\|^2.$$

Пользуясь этой оценкой при $\varepsilon=2$ в (27), придём к неравенству

$$w_k \le a_7\left\|\mathbf{v}^{k-1}-\mathbf{v}^{k}\right\|^2-a_8\left\|\mathbf{y}^{k}\right\|^2-a_9\left\|\mathbf{x}^{k}-\mathbf{v}^{k-1}\right\|^2,\quad k\ge 0, \tag{28}$$

где $a_8=a_{21}(1-s_k)-a_4-3a_6(1+s_k);\quad a_9=1-s_k-(a_3+1.5a_6)(1+s_k).$

Заметим, что существует число $N>1$ такое, что $\forall\ \ k-1>N\ge k_0$ с учетом (9) наряду с (25') выполняются неравенства:

$$a_6(1+s_k)\le 4/15-7\alpha/5;\quad a_{11}\le a_1<2/5;$$

$$a_{21}(1-s_k)\ge 1-2\alpha-s_k;\quad a_3(1+s_k)\le 3/5+2\alpha-s_k; \tag{29}$$

$$a_3\le 3/5+2\alpha;\quad a_4\le 4\alpha-s_k;\quad 0<\alpha<1/9;\quad 0<s_k<1.$$

С учетом (9), (25') и (29), в (28) для коэффициентов имеют место оценки:

$$a_7 \le \left(a_3 + 2a_{11}\right)(1 + s_k^{-1}) \le (7/5 + 2\alpha)(1 + s_k^{-1});$$

$$a_8 \ge 1 - 2\alpha - s_k - 4\alpha + s_k - 3(4/15 - 7\alpha/5) = 1/5 - 9\alpha/5 > 0;$$

$$a_9 \ge 1 - s_k - 3/5 - 2\alpha + s_k - 1.5(4/15 - 7\alpha/5) > 0.$$

Учитывая эти оценки и (13), из (28) и (25) получим:

$$w_k \le a_7 \left\| \mathbf{v}^{k-1} - \mathbf{v}^k \right\|^2 \le (7/5 + 2\alpha) C^2 (1 + s_k^{-1}) \tau_k^{-2} \left| \tau_{k+1} - \tau_k \right|^2;$$

$$u_k \le (1 - s_k) u_{k-1} + d_k, \quad k - 1 \ge N, \tag{30}$$

где

$$d_k = w_k + a_5 = C^2 (7/5 + 2\alpha)(1 + s_k^{-1}) \tau_k^{-2} \left| \tau_{k+1} - \tau_k \right|^2 + C_1 \beta^2 (\delta + \tau).$$

Здесь, ввиду (9) и (29), имеют место соотношения

$$\sum_{k=0}^{k=\infty} s_k = \infty, \quad \frac{d_k}{s_k} = \frac{w_k + a_5}{s_k} \to 0, \quad k \to \infty.$$

Последовательность $\{u_k\}$, обладающая такими свойствами, удовлетворяющая (25), (30), стремится к нулю (см., например, [14], с. 96). Сходимость $\{u_k\}$ равномерная относительно выбора приближений $\nabla \mathbf{f}_k(\mathbf{x})$ в (7), т.к. s_k и d_k в (30) не зависят от этих приближений. Отсюда следует равномерная сходимость в (15). Из (15), (11) и (14) следует утверждение теоремы 1. ♦

Примечание. В качестве параметров метода (5), (6), удовлетворяющих условиям теоремы 1, могут быть выбраны, например, следующие:

$$\alpha_k = c_1 (1 + k)^{-1}, \quad \beta_k = c_2 (1 + k)^{-1},$$

$$\tau_k = c_3 (1 + k)^{-1}, \quad \delta_k = c_4 (1 + k)^{-5/2}, \tag{31}$$

где числа $c_i > 0, \quad i \in 1{:}4; \; c_1 \ge c_2 > c_3 > c_4$.

5. **Правило останова метода**

Правило останова метода (5), (6), (9) строится аналогично тому, как это сделано, например, в работах [6]–[8]. Согласно теореме 1, градиент $\nabla f(\mathbf{x})$ функции $f(\mathbf{x})$ может быть вычислен в каждой фиксированной точке $\mathbf{x}^0 \in Q$ в задаче (1) с некоторой ошибкой δ_k, такой, что $\delta_k \to 0$ при $k \to \infty$, и выполняется (7). Однако, в конкретных задачах ошибка начальных данных не обязательно стремится к нулю, а обычно больше некоторого фиксированного положительного числа. Предположим, что для каждого фиксированного $\mathbf{x} \in H$, вместо вычисления точного значения градиента $\nabla f(\mathbf{x})$, для заданного числа w, возможно вычислить его аппроксимацию $\nabla f_w(\mathbf{x})$ такую, что

$$\max_{\mathbf{x} \in H} \left\|\nabla f_w(\mathbf{x}) - \nabla f(\mathbf{x})\right\| \le w(1+\|\mathbf{x}\|), \qquad (32)$$

где $w > 0$ известное число. Тогда, заменяя приближённый градиент $\nabla f_k(\mathbf{x})$ в методе (5),(6),(9) на $\nabla f_w(\mathbf{x})$ из (32), вместо (5),(6) приходим к методу

$$\mathbf{y}^k = \mathbf{x}^k - \mathbf{x}^{k-1}; \quad \mathbf{z}^k = \mathbf{x}^k + \alpha_k \mathbf{y}^k;$$

$$\mathbf{x}^{k+1} = P_Q\left\{\mathbf{z}^k - \beta_k \mathbf{B}(\mathbf{z}^k)^{-1}\left[\nabla f_w(\mathbf{z}^k) + \tau_k \mathbf{z}^k\right]\right\}, \quad k = 0,1,2,\ldots \qquad (33)$$

Однако, для метода (33) условия (9) согласования параметров $\alpha_k, \beta_k, \tau_k$ с параметром погрешности $\delta_k \equiv w > 0, \quad k \ge 0$, будут очевидно нарушены, так что процесс (33) может расходиться и его использование для больших значений k будет неверно. Опираясь на теорему 1, можно построить правило останова и получить ответ на важный вопрос: до какого номера итерации $k \ge 1$ нужно продолжить процесс (33), чтобы полученную в его результате точку $\mathbf{x}^{k+1}$ можно было взять за приближение к нормальному решению $\mathbf{x}^*$ задачи (1), соответствующее выбранному уровню погрешности $w>0$. Для этого фиксируем некоторую начальную точку $\mathbf{x}^0 \in H$ и последовательности параметров метода

$$\{\alpha_k\}, \{\beta_k\}, \{\tau_k\}, \{\delta_k\}, \tag{34}$$

удовлетворяющие условиям (9). Могут быть выбраны параметры (31) и $w<\delta_0$. Поскольку выполнение условий (9) здесь не предполагается, то $w=w_k$, $k \ge 0$, теперь является параметром метода (33), совершенно не связанным с условиями (9). Для каждого фиксированного $w \in [0;\delta_0]$ в (32), процесс (33) с выбранными параметрами (34) будем продолжать до номера итерации $k_0 \ge 1$, определённого из условия

$$\delta_k \ge w, \quad k = 0,1, \dots, k(w). \tag{35}$$

Так как $\delta_k \to +0$ при $k \to \infty$, $\delta_0 > w$, то требуемый номер итерации $k(w) \ge 1$ будет конечным и найдется для каждого фиксированного $w>0$. В теореме 2 дается обоснование критерия (35) останова процесса (33).

Теорема 2. *Пусть:* 1) *выполнены все условия теоремы 1;* 2) *приближения* $\nabla f_w(\mathbf{x})$ *градиентов* $\nabla f(\mathbf{x})$ *удовлетворяют условию* (32)*;*

3) *множество точек* $\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^{k(w)}$ *получено методом* (33)*, где номер итерации* $k(w) \ge 1$ *определён в соответствии с правилом останова* (35)*. Тогда*

$$\left\|\mathbf{x}^{k(w)} - \mathbf{x}^*\right\| \to 0 \ \text{ при } \ w \to +0, \tag{36}$$

в том смысле, что для сколь угодно малых $w>0, \varepsilon>0$ *найдется номер* $k(w)$ *такой, что* $\left\|\mathbf{x}^{k(w)} - \mathbf{x}^*\right\| < \varepsilon$.

Доказательство проводится аналогично соответствующим теоремам из работ [6]-[11]. В условиях теоремы 2 из (32) и правила останова (35) следует,

$$max\left\{\left\|\nabla f_w(\mathbf{x}) - \nabla f(\mathbf{x})\right\|\right\} \le \delta_k (1 + \|\mathbf{x}\|), \ \mathbf{x} \in H,$$

$k = 0,1,2,...,k(w)$. Тогда градиенты $\nabla f_w(\mathbf{x})$, удовлетворяют условию (7) $\forall$ $k = 0,1,2,...,k(w)$. По правилу останова номер $k(w)$ из (35) обеспечивает выполнение условия $w < \delta_0$. Поэтому при $\delta_k \to 0$ имеем $k(w) \to \infty$, $w \to +0$, т.е. останов метода будет осуществляться при очень больших номерах $k(w)$. По теореме 1 последовательность $\{\mathbf{x}^k\}$ метода (33) сходится по норме H к решению $\mathbf{x}^*$ задачи (1), поэтому $\forall$ $\varepsilon > 0$ существует номер $k(\varepsilon)$, не зависящий от приближенных градиентов $\nabla f_k(\mathbf{x})$ в (7) и $\mathbf{x}^{k+1}$ из (33), что

$$\left\|\mathbf{x}^k - \mathbf{x}^*\right\| < \varepsilon \quad \forall \ k > k(\varepsilon). \tag{36'}$$

виду соотношения $k(w) \to +\infty$ существует число $w(\varepsilon) > 0$ такое, что $k(w(\varepsilon)) > k(\varepsilon)$ $\forall$ $w \in (0, w(\varepsilon))$. Поэтому точки $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^{k(w)}$ соответствуют точкам метода (5), (6) при $\nabla f_k(\mathbf{x}) = \nabla f_w(\mathbf{x})$, $k = 0,1,2,...,k(w)$, и условиям (9), с учетом (36'). Поскольку $k(w) > k(\varepsilon)$, из (36') следует неравенство $\left\|\mathbf{x}^{k(w)} - \mathbf{x}^*\right\| < \varepsilon$, верное $\forall$ $\varepsilon > 0$, при $\forall$ $w \to +0$, поэтому оно влечет (36). Теорема 2 доказана.

6. **Регуляризующий оператор**

Регуляризующий оператор (см. [1]–[5]) задачи (1) строим способом, аналогичным использованным в работах [4]–[10]. Предположим, что погрешность $w > 0$ фиксирована и выполнено правило останова (35). Соотношение (36) соответствует правилу останова (35) и каждому фиксированному уровню ошибки $w > 0$ в (32). При таком уровне погрешности в (32), выполнении правила останова (35) и теоремы 2, регуляризующим будет оператор $\mathbf{R}_w = \mathbf{R}_w\left(\nabla \mathbf{f}_w\left(\mathbf{x}^{k(w)}\right), w\right)$, сопоставляющий всякому набору своих аргументов, т.е. числу $w \in (0, w(\varepsilon)]$ (где $\varepsilon -$ заданная точность вычисления минимума) и приближению $\nabla f_w(\mathbf{x})$ градиента функции $f(\mathbf{x})$ из (32), точку $\mathbf{x}^{k(w)}$ метода (33),(35). Входящие в определение регуляризующего оператора $\mathbf{R}_w$, $0 < w \le w(\varepsilon)$, параметры метода (33) удовлетворяют условиям (9), (32), а в остальном произвольны.

7. **Оценка скорости сходимости метода**

Оценку скорости сходимости метода (5) получим при более строгих предположениях, чем при доказательстве сходимости в теореме 1.

Теорема 3. *Пусть выполнены все условия теоремы 1 и, кроме того:* 1) *функция* $\varphi(\mathbf{x}) \in C^{1,1}(Q)$ *сильно выпуклая с константой сильной выпуклости* $\kappa > 0$;

2) *параметры* $\alpha_k, \beta_k, \tau_k, \delta_k$ *метода* (5), (6) *таковы, что*

$$\alpha_k \to \alpha^0 > 0, \;\; \alpha^0 \le \alpha_k \le 1/50, \quad \tau_k \to 0, \;\; k \to \infty;$$

$$\beta \le min\left\{\frac{m\alpha - 2\tau - \delta}{\delta + (1+\alpha)\tau}; \frac{28\alpha}{(L+\mu)(1+\alpha)}; \frac{5(L+2\kappa)\alpha}{4L\kappa};\right.$$

$$\left.\frac{4M(3m - 5\delta - 10\tau)}{5\left\{M\left[4\delta + (1+\alpha)\left[4\tau + m(L+\mu)\right]\right] - 8m(1-\alpha)\tau\right\}}\right\}. \qquad (37)$$

Тогда последовательность $\{\mathbf{x}^k\}$, *определяемая методом* (5), (6), (9), (37), *из произвольной начальной точки* $\mathbf{x}^0 \in Q$, *равномерно относительно выбора приближённых градиентов из* (7), *сходится к решению* $\mathbf{x}^*$ *задачи* (1) *с оценкой*

$$\|\mathbf{x}^{k+1} - \mathbf{x}^*\| \le \left[r_k \|\mathbf{x}^k - \mathbf{x}^*\|^2 + e_k\right]^{1/2}, \qquad (38)$$

где $0 < r_k = qp < 1$, $q = 1 - 2L\kappa\beta_k/(L+2\kappa)$, $e_k = b_5 p$, $p = \frac{6-30\alpha_k}{1-48\alpha_k}$; *множитель* $b_5 = C_1\beta^2(\delta+\tau)$; $C_1 = max\left\{(C+1)^2/m; 4C^2/m\right\}$.

Доказательство. Прежде всего заметим, что в условиях теоремы метод сходится $\forall\; \mathbf{x}^0 \in Q$ равномерно относительно выбора приближённых градиентов из (7). Выкладками, аналогичными проведенным в теореме 1 при получении неравенства (24), пользуясь для скалярного произведения оценкой (см. [14], с. 188)

$$(\nabla\varphi(\mathbf{u}) - \nabla\varphi(\mathbf{v}), \mathbf{v} - \mathbf{w}) \le (L+\mu)\|\mathbf{u} - \mathbf{w}\|^2/4 - L\mu\|\mathbf{u} - \mathbf{v}\|^2/(L+\mu),$$

$\mathbf{u}, \mathbf{v}, \mathbf{w} \in Q$, справедливой для сильно выпуклой функции $\varphi(\mathbf{x}) \in C^{1,1}(Q)$, а также (18), (21), при $\mathbf{x}^* \in Q_*$, от (20) придём к неравенству

$$\left\|\mathbf{x}^{k+1} - \mathbf{x}^*\right\|^2 - b_1 \left\|\mathbf{x}^k - \mathbf{x}^*\right\|^2 + b_2 \left\|\mathbf{x}^{k+1} - \mathbf{x}^k\right\|^2 \le$$

$$\le b_3 \left\|\mathbf{x}^k - \mathbf{x}^*\right\|^2 + b_4 \left\|\mathbf{y}^k\right\|^2 - \alpha \left\|\mathbf{x}^{k-1} - \mathbf{v}^k\right\|^2 + b_5, \quad k \ge 0, \tag{39}$$

где $b_1 = \frac{2}{5} - \alpha - 2(1-\alpha)\beta\tau / M;\ b_2 = 1 - \alpha - \frac{\beta}{2m}[Lm + \alpha(Lm + 2\tau) + 4\delta + 2\tau] -$

$$-(2\tau + \delta)/m; \quad b_3 = \frac{3}{5} - \frac{2L\kappa(1-\alpha)\beta}{L + 2\kappa} + \frac{2(\tau + \delta)}{m} + \frac{\beta[(3+\alpha)\delta + (1+\alpha)\tau]}{m};$$

$$b_4 = 2\alpha + \left(\alpha - \alpha^2\right)\left(\frac{2}{M} + \frac{2L\kappa}{L + 2\kappa}\right)\beta + \left(\alpha^2 + \alpha\right)\beta\left(\frac{\delta}{m} + \frac{2\tau}{m} + \frac{L + 2\kappa}{4}\right).$$

Представим неравенство (39) в виде

$$u_{k+1} \le (1 - s_k)u_k + v_k + b_5, \quad k \ge 0, \tag{40}$$

где $s_k = s = 2L\kappa\beta_k(L + 2\kappa)$, левая часть u_{k+1} равна левой части (39), а v_k равно разности между правой частью (39) без b_5 и первым слагаемым в правой части (40):

$$u_{k+1} = \left\|\mathbf{x}^{k+1} - \mathbf{x}^*\right\|^2 - b_1 \left\|\mathbf{x}^k - \mathbf{x}^*\right\|^2 + b_2 \left\|\mathbf{x}^{k+1} - \mathbf{x}^k\right\|^2;$$

$$v_k = b_3 \left\|\mathbf{x}^k - \mathbf{x}^*\right\|^2 - \alpha \left\|\mathbf{x}^{k-1} - \mathbf{x}^*\right\|^2 + [b_4 - (1-s)b_{21}]\left\|\mathbf{y}^k\right\|^2 - (1-s)\left\|\mathbf{x}^k - \mathbf{x}^*\right\|^2 +$$

$$+ (1-s)b_{11} \left\|\mathbf{x}^{k-1} - \mathbf{x}^*\right\|^2 = b_6 \left\|\mathbf{x}^{k-1} - \mathbf{x}^*\right\|^2 - b_7 \left\|\mathbf{x}^k - \mathbf{x}^*\right\|^2 - b_8 \left\|\mathbf{x}^k - \mathbf{x}^{k-1}\right\|^2, k \ge 0,$$

где

$$b_6 = (1-s)b_{11} - \alpha,\ b_7 = 1 - b_3 - s,\ b_8 = (1-s)b_{21} - b_4,\ b_{11} = b_{1(k-1)},\ b_{21} = b_{2(k-1)}.$$

Оценим левую часть (40) снизу, а v_k в правой части (40) – сверху. Для этого заметим, что при условиях теоремы 3 имеют место следующие неравенства:

$$1/5 < b_1 < 2/5;\ 9/10 - 4\alpha \le b_2 \le 1 - 3\alpha;\ b_3 \le 1/5 + \delta; \quad b_1 s < 2\alpha;$$

$$b_4 \le 9\alpha/2;\ 1/5-3\alpha < b_6 < 2/5-2\alpha;\ b_2 - b_1 - b_1 b_2 \ge 0.1 - 24\alpha/5 > 0;$$

$$b_7 > 4/5 - 3\alpha - \delta > 0,\ b_8 \ge 0.9 - 9\alpha > 0. \qquad (41)$$

Пользуясь (41) и следующим из (21) неравенством

$$\left\|\mathbf{x}^k - \mathbf{x}^*\right\|^2 \le (1+\varepsilon)\left\|\mathbf{x}^k - \mathbf{x}^{k+1}\right\|^2 + (1+\varepsilon^{-1})\left\|\mathbf{x}^{k+1} - \mathbf{x}^*\right\|^2$$

при $\varepsilon = (b_2 - b_1)/b_1$, получим оценку снизу

$$u_{k+1} \ge [1 - b_1 b_2/(b_2 - b_1)]\left\|\mathbf{x}^{k+1} - \mathbf{x}^*\right\|^2, \qquad (42)$$

где при условиях (41) имеем $1 - b_1 b_2/(b_2 - b_1) \ge \gamma = (1-48\alpha)/(6-30\alpha)$.

Оценивая v_k в (40) сверху, с учетом (41) и неравенства

$$\left\|\mathbf{x}^{k-1} - \mathbf{x}^*\right\|^2 \le 2\left\|\mathbf{x}^{k-1} - \mathbf{x}^k\right\|^2 + 2\left\|\mathbf{x}^k - \mathbf{x}^*\right\|^2,$$

получаем

$$v_k \le (2b_6 - b_7)\left\|\mathbf{x}^k - \mathbf{x}^*\right\|^2 - (b_8 - 2b_6)\left\|\mathbf{y}^k\right\|^2, \qquad (43)$$

где с учётом (41) $2b_6 - b_7 \le -\delta - \alpha < 0,\quad b_8 - 2b_6 \ge 1/10 - 5\alpha \ge 0$. Отсюда имеем $v_k < 0$ и из (40), (42), (43) следует оценка

$$\left\|\mathbf{x}^{k+1} - \mathbf{x}^*\right\|^2 \le r_k \left\|\mathbf{x}^k - \mathbf{x}^*\right\|^2 + e_k, \qquad (44)$$

где, с учетом (9), (37), (41), $0 < r_k = (1-s)\gamma < 1;\quad e_k = b_5\gamma$. Из (44) следует (38). Теорема 3 доказана.

8. Заключение

Полученные результаты по исследованию проекционных экстраградиентных обобщённых двухточечных двухэтапных квазиньютоновских методов в основном завершают теорию этих методов. Возможно появление новых версий ПОДЭМК, но обоснования сходимости и оценок скорости сходимости для них не будут значительно отличаться от полученных (скорости сходимости от линейной до квадратичной для нерегуляризованных базовых методов). Работоспособность методов проверена с помощью сравнительных численных экспериментов с другими известными методами минимизации первого и второго порядка на известных тестовых задачах минимизации. Проверка показывает перспективность проекционных обобщённых двухточечных экстраградиентных методов переменной метрики. Эти методы способны с высокой точностью решать такие задачи минимизации, с которыми другие методы переменной метрики и методы второго порядка не справляются.

Литература

1. *Тихонов А.Н.* О некорректно поставленных задачах // Вычислит. математ. и прогр. Вып. 8. Изд-во МГУ, 1967. С. 3-33.

2. *Антипин А.С.* Об едином подходе к методам решения некорректных экстре-мальных задач // Вестник МГУ. Сер. 1. Математика и механика. 1973. № 2. С. 60-67.

3. *Васильев Ф.П.* Методы решения экстремальных задач. М.: Наука, 1981. 400 с.

4. *Васильев Ф.П.* Методы решения неустойчивых экстремальных задач с неточно заданными исходными данными. Дисс. ... д.ф.-м.н. М., 1985. 340 с.

5. *Бакушинский А.Б., Гончарский А.В*. Итеративные методы решения некорректных задач. М.: Наука, 1989. 128 с.

6. *Васильев Ф.П., Недич А.* Об одном варианте регуляризованного метода проекции градиента // Журнал вычисл. матем. и матем. физики. 1994. Т. 34. № 4. С. 511-519.

7. *Васильев Ф.П., Амочкина Т.В., Недич А.* Об одном регуляризованном варианте двухшагового метода проекции градиента // Вестник МГУ. Сер. 15. Вычисл. матем. и киберн. 1996. № 1. С. 35-42.

8. *Васильев Ф.П., Недич А.* О трёхшаговом регуляризованном методе проекции градиента для решения задач минимизации с неточными исходными данными // Изв.вузов. Математика. 1993. № 12. С. 35-43.

9. *Васильев Ф.П., Недич А.* О четырёхшаговом регуляризованном методе проекции градиента для задач минимизации с неточными исходными данными // Mathematica montesnigri. 1995. V. 4. P. 83-101.

10. *Малинов В.Г.* Четырёхпараметрический двухшаговый регуляризованный проекционный метод минимизации // Журнал вычисл.матем.и матем. физ. 1999. Т. 39. № 4. С. 567-572.

11. *Васильев Ф.П., Недич А.* Регуляризованный непрерывный метод проекции градиента второго порядка // Вестник МГУ. Сер. 15. Вычисл. матем. и киберн. 1994. № 2. С. 3-11.

12. *Малинов В.Г.* Регуляризованный проекционный непрерывный метод для задач минимизации с ограничениями // ЖВМиМФ. 2002. Т. 42. № 11. С. 1646-1656.

13. *Рязанцева И.П.* Об одном методе итеративной регуляризации для выпуклых задач минимизации // ЖВМиМФ. 2000. Т. 40. № 2. С. 181–187.

14. *Малинов В.Г.* Непрерывные методы минимизации второго порядка // Учёные записки ОО ИПК РО. 2003. Т. 31. С. 102-131.

15. *Антипин А.С., Васильев Ф.П.* О непрерывном методе минимизации в пространствах с переменной метрикой // Известия вузов. Математика. 1995. № 12(403). С. 3–9.

16. *Амочкина Т.В., Антипин А.С., Васильев Ф.П.* Регуляризованный непрерывный метод минимизации с переменной метрикой при неточно заданных исходных данных // Вестник МГУ. Сер. 15. Вычисл. матем. и киберн. 1996. № 4. С. 5–11.

17. *Амочкина Т.В.* Непрерывный метод проекции градиента второго порядка с переменной метрикой // Журнал вычисл. матем. и матем. физики. 1997. Т. 37. № 10. С. 1174–1182.

18. *Malinov V.G.* On Continuous Projection Minimization Method of the Second Order with a Variable Metric // The International Conference on Applied Mathematics Dedicated to the 65-th Anniversary of B.N. Pshenichnyi (1937-2000). June 25-28. Abstracts. Ukraine. Kiev. NTUU, 2002. P. 48–49.

19. *Malinov V.G.* Regularized continuous projection minimization method of the second order in variable metric space // International Conference "Ill-posed and inverse problems" dedicated to prof. M.M. Lavrent'ev... August 5-9, 2002. Abstracts. Novosibirsk. Sobolev Inst. of Mathematics, 2002. P. 113.

20. *Малинов В.Г.* Проекционный двухшаговый обобщенный двухпараметрический метод минимизации первого порядка с переменной метрикой // Ученые записки Ульяновского гос.университета. Серия "Фундаментальные проблемы математики и механики". Вып.1(13). Ульяновск: УлГУ, 2003. С. 127–138.

21. *Малинов В.Г.* Обобщенный двухпараметрический двухшаговый проекционный метод минимизации второго порядка с переменной метрикой // Прикладная математика и механика. Сборник научных трудов. Ульяновск: УлГТУ, 2004. С. 121–129.

22. *Васильев Ф.П.* Численные методы решения экстремальных задач. М.: Наука, 1988. 552 с.

23. *Malinov V.G.* On regularized two-step variable metric minimization method // V Moscow International conference on operations research (ORM2007) dedicated to the outstanding Russian scientist Nicita N. Moiseev 90th birthday. Moscow, April 20-14. Proceedings. M.: MAKS Press, 2007. P. 171-173.

24. *Антипин А.С.* Методы нелинейного программирования, основанные на прямой и двойственной модификации функции Лагранжа. Препринт ВНИИ системных исследований. М., 1979. 74 с.

25. *Le D.* A Fast and Robust unconstrained optimization method requiring minimum storage // Mathematical Programming. 1985. V. 32. № 1. P. 41-68.

26. *Wolfe M.A.* A Quasi-Newton Method with Memory for Unconstrained Function Minimization // J. Inst. Mathematics & Applications. 1975. V. 15. N. 1. P. 85-94.

27. *Колмогоров А.Н.*, *Фомин С.В.* Элементы теории функций и функционального анализа. М.: Физматлит, 1976. 544 с.