УДК 519.853.62

ПОДМ с проектированием в переменной метрике

© В. Г. Малинов [1]

Аннотация. В работе впервые исследуется проекционный обобщённый двухшаговый двухэтапный метод (ПОДМ) с проектированием в переменной метрике (ПОДМПМ) для решения конечномерных задач минимизации на выпуклом замкнутом множестве в евклидовом пространстве, для решения функциональных уравнений и других задач. Сходимость метода доказана для выпуклых гладких функций с Липшицевыми градиентами. Получены оценки скорости сходимости: линейной - для выпуклых гладких функций, сверхлинейной и квадратичной - для дважды гладких функций при дополнительных условиях.

Ключевые слова: ПОДМПМ, проектирование в переменной метрике, сходимость, скорость сходимости.

1. Постановка задачи

Рассмотрим задачу минимизации на выпуклом замкнутом простом множестве

$$f(\mathbf{x}) \longrightarrow \inf, \quad \mathbf{x} \in Q \subset E^n, \tag{1.1}$$

где n-мерное евклидово пространство E^n нормировано скалярным произведением, $\|\mathbf{x}\| = (\mathbf{x},\mathbf{x})^{1/2}\ \forall \mathbf{x} \in E^n$, выпуклая функция $f(\mathbf{x}) \in C^{1,1}(Q)$. Предполагаем, что функция $f(\mathbf{x})$ имеет *гиперповерхности уровней овражной структуры* и ограничена снизу, множество её минимумов не пусто:

$$\inf f(\mathbf{x}) = f_* > -\infty,\ \mathbf{x} \in Q; \quad Q_* = \{\mathbf{x} \in Q :\ f(\mathbf{x}) = f_*\} \neq \emptyset. \tag{1.2}$$

Множество сложных задач поставленного вида не решаются существующими методами или методы имеют низкую скорость сходимости в окрестности минимума функции, ввиду наличия овражности функции. Традиционные, ставшие классическими, многошаговые методы проекции градиента (МПГ) (см., например, [1]– [4]) без специальной процедуры локального поиска этим недостатком также обладают. В связи с этим предложены непрерывные МПГ в пространствах с переменной метрикой (НМПГПМ) [5]– [6] для задач, описываемых обыкновенными дифференциальными уравнениями. Для других задач оптимизации вида (1.1), не связанных с непрерывными моделями, предложены и исследованы итеративные методы с переменной метрикой класса ПОДМ [7]– [8]. Кратко напомним, что *проекционные обобщённые двухшаговые двухэтапные* методы (ПОДМ) минимизации функций с «овражными» гиперповерхностями уровней — это класс проекционных двухшаговых методов, строящих минимизирующую последовательность $\{\mathbf{x}^k\} \longrightarrow \mathbf{x}^* \in Q_*$ с помощью прогнозной точки $\mathbf{z}^k = \mathbf{x}^k + \alpha_k \mathbf{y}^k$ на «склоне оврага» (где $\mathbf{y}^k = \mathbf{x}^k - \mathbf{x}^{k-1}$) или проекции $\mathbf{z}^k = P_Q\left(\mathbf{x}^k + \alpha_k \mathbf{y}^k\right)$ (где α_k – один из параметров методов класса) и проводящие минимизацию из точки $\mathbf{z}^k$ вдоль направлении убывания функции — линейной комбинации вектора $\mathbf{y}^k$ и антиградиента $-\nabla f(\mathbf{z}^k)$ (см., например, [4], [7], [8]). Ввиду использования градиентов $\nabla f(\mathbf{z}^k)$ большей величины на склоне оврага (в отличие от методов, использующих малый градиент $\nabla f(\mathbf{x}^k)$ на «дне» оврага), ПОДМ мало чувствительны к овражности функций и ошибкам округлений [2].

[1] Доцент кафедры ЭММиИТ, Ульяновский госуниверситет, г. Ульяновск; vgmalinov@mail.ru.

Исследуемые ПОДМ с проектированием в переменной метрике обладают свойствами, объединяющими достоинства методов класса ПОДМ и методов переменной метрики (последние для задач безусловной минимизации называют квазиньютоновскими).

Отметим, что: 1) здесь идея, реализованная при построении НМПГПМ [5], распространена на итеративные ПОДМ квазиньютоновского типа; 2) основанные на другой идее многошаговые методы переменной метрики для *решения задач безусловной минимизации* построены иначе (см., например, [9]– [10]); 3) в данной работе полные результаты исследования ПОДМ с проектированием в переменной метрике излагаются впервые; 4) здесь получены оценки линейной, сверхлинейной и квадратичной скорости сходимости ПОДМПМ.

2. Пространство E_1^n и предлагаемый метод

Наряду с существующей метрикой и оператором $P_Q(\mathbf{v})$ проектирования вектора $\mathbf{v} \in E^n$ на множество Q, введём в E^n новую метрику с помощью нового скалярного произведения $(\mathbf{B}(\mathbf{x})\mathbf{u},\mathbf{u})\ \forall\,\mathbf{u},\mathbf{x} \in E^n$, где $\mathbf{B}(\mathbf{x}):\ E^n \longrightarrow E^n$, при каждом фиксированном $\mathbf{x} \in E^n$, суть положительно определённый самосопряженный линейный оператор метрики пространства; $P_Q^{\mathbf{B}(\mathbf{x})}[\mathbf{v}]$ суть оператор проектирования в метрике $\mathbf{B}(\mathbf{x})$ вектора $\mathbf{v} \in E^n$ на множество Q. Критерием проекции $\mathbf{w} = P_Q^{\mathbf{B}(\mathbf{x})}[\mathbf{v}] \in Q$ в новой метрике служит неравенство [6]

$$(\mathbf{B}(\mathbf{x})(\mathbf{w}-\mathbf{v}),\mathbf{u}-\mathbf{w}) \geq 0,\ \mathbf{u} \in Q. \tag{2.1}$$

Проекция существует и единственна как решение $\mathbf{w} \in Q$ квадратичной задачи

$$g(\mathbf{u}) = (\mathbf{B}(\mathbf{x})(\mathbf{u}-\mathbf{v}),\mathbf{u}-\mathbf{v}) \longrightarrow \inf,\ \mathbf{u} \in Q, \tag{2.2}$$

в силу выпуклости множества Q и сильной выпуклости функции $g(\mathbf{u})$ [6]. Полученное евклидово пространство с двумя скалярными произведениями и определяемыми ими метриками обозначим E_1^n, далее подразумеваем задачу вида (1.1) в нём. В построенном пространстве для решения задачи исследуем ПОДМПМ:

$$1\ \text{этап.}\mathbf{y}^k = \mathbf{x}^k - \mathbf{x}^{k-1},\ \mathbf{z}^k = \mathbf{x}^k + \alpha_k\mathbf{y}^k,\ \mathbf{x}^0,\mathbf{x}^1 \in E_1^n,\ f(\mathbf{x}^0) > f(\mathbf{x}^1); \tag{2.3}$$

$$2\ \text{этап.}\quad \mathbf{x}^{k+1} = P_Q^{\mathbf{B}(\mathbf{z}^k)}\left[\mathbf{z}^k - \beta_k[\mathbf{B}(\mathbf{z}^k)]^{-1}\nabla f(\mathbf{z}^k)\right], k = 1,2,..., \tag{2.4}$$

где $\mathbf{x}^0$ – начальная точка; точку $\mathbf{x}^1 \neq \mathbf{x}^0$ выбираем такую, что $f(\mathbf{x}^0) > f(\mathbf{x}^1)$; α_k,β_k – параметры метода; $\mathbf{B}(\mathbf{z}^k) = \mathbf{B}_k$ – последовательность положительно определённых самосопряжённых линейных операторов. В зависимости от выбора операторов $\mathbf{B}_k$ и способов выбора параметров метода, из (2.3), (2.4) получаем различные ПОДМПМ первого порядка, а при $\mathbf{B}(\mathbf{z}^k) = \nabla^2 f(\mathbf{z}^k)$ — ПОДМ второго порядка. Оператор $\mathbf{B}(\mathbf{x})$ в (2.4) таков, что

$$m\|\mathbf{u}\|^2 \leq (\mathbf{B}(\mathbf{x})\mathbf{u},\mathbf{u}),\quad m > 0,\ \forall\,\mathbf{u},\mathbf{x} \in E_1^n. \tag{2.5}$$

Отметим, что в отличие от методов из работ [7], [13], ввиду использования в (2.3), (2.4) проекции $P_Q^{\mathbf{B}(\mathbf{x})}(\mathbf{v})$ в метрике $\mathbf{B}(\mathbf{x})$, а не проекции в обычной метрике $P_Q(\mathbf{v})$, возникают трудности в доказательстве сходимости, которые преодолеваются с помощью получаемого в лемме 5 нового неравенства. При вычислительной реализации метода (2.3), (2.4) задача (2.2) решается численно.

К операторам $\mathbf{B}(\mathbf{x})$ с указанными выше свойствами, кроме выполнения (2.5), в работе предъявляется требование выполнения неравенства $(\mathbf{B}^{-1}(\mathbf{x}^*)\nabla f(\mathbf{x}^*),\mathbf{u}-\mathbf{x}^*) \geq 0\ \forall\ \mathbf{u} \in$

$Q, \mathbf{x}^* \in Q_*$. Это неравенство доказывается в леммах 1 и 2. Предполагается, что формируемые в (2.4) векторы $\mathbf{p}(\mathbf{z}^k) = -[\mathbf{B}(\mathbf{z}^k)]^{-1}\nabla f(\mathbf{z}^k)$, $k = 1, 2, \ldots$ задают направления убывания функции $f(\mathbf{x})$ в точках $\mathbf{z}^k$ и составляют острый угол с направлением антиградиента $-\nabla f(\mathbf{z}^k)$, направлены в сторону точки минимума функции; следовательно, $(\nabla f(\mathbf{z}^k), \mathbf{p}^k) > 0$, *это не оговаривается в дальнейшем.* В теореме 4 используется сходимость последовательности операторов $\mathbf{B}_k$ к гессиану. Обобщение неравенства из работы [11] для $\mathbf{B}_k$ применяется в доказательстве теоремы 5.

3. Вспомогательные утверждения

Для независимости изложения от других источников, приведём вспомогательные утверждения для доказательства сходимости и оценок скорости сходимости метода семейства (2.3), (2.4). Большая часть из них доказана в работах [7], [13].

Примечание 1. По самому построению пространства E_1^n в нём, наряду с (2.1), имеет место критерий ([1], с. 189)

$$(\mathbf{w} - \mathbf{v}, \mathbf{u} - \mathbf{w}) \geq 0 \ \forall \mathbf{u} \in Q \tag{3.0}$$

проекции $\mathbf{w} \in Q$ вектора $\mathbf{v} \in E_1^n$ на выпуклое замкнутое множество $Q \subset E_1^n$. Поскольку евклидовы пространства с различными скалярными произведениями изоморфны ([12], с. 124, 127), в E_1^n сохраняются все соотношения и теоремы из E^n, связанные со скалярным произведением $(\mathbf{x}, \mathbf{x})$.

В леммах даны соотношения, полезные при доказательстве сходимости и оценке скорости сходимости ПОДМПМ и других методов класса ПОДМ в E^n и E_1^n, для определённости они доказываются в одном из этих пространств.

Лемма 1. *Пусть выпуклые функции $f(\mathbf{x})$ и $\varphi(\mathbf{x})$ класса $C^{1,1}(E_1^n)$ таковы, что*

$$\nabla\varphi(\mathbf{x}) = \mathbf{B}^{-1}(\mathbf{x})\nabla f(\mathbf{x}) \ \forall \ \mathbf{x} \in E_1^n \tag{3.1}$$

и множество точек минимума функций $f(\mathbf{x})$ и $\varphi(\mathbf{x})$ не пусто, $Q_* \neq \emptyset$.

Тогда для $\mathbf{x}^* \in Q_*$ в пространстве E_1^n имеет место неравенство

$$\left(\mathbf{B}^{-1}(\mathbf{x}^*)\nabla f(\mathbf{x}^*), \mathbf{u} - \mathbf{x}^*\right) \geq 0 \ \forall \ \mathbf{u} \in Q. \tag{3.2}$$

Д о к а з а т е л ь с т в о. Пользуясь формулой Лагранжа для функции $\varphi(\mathbf{x})$, определением точки минимума $\mathbf{x}^* \in Q_* \subset Q \subset E_1^n$ и (3.1), получим:

$$\varphi(\mathbf{u}) - \varphi(\mathbf{x}^*) = (\nabla\varphi(\mathbf{x}^*), \mathbf{u} - \mathbf{x}^*) = \left(\mathbf{B}^{-1}(\mathbf{x}^*)\nabla f(\mathbf{x}^*), \mathbf{u} - \mathbf{x}^*\right) \geq 0, \ \mathbf{u} \in Q.$$

Лемма 1 доказана.

Лемма 2. *Пусть множество $Q \subset E_1^n$ выпукло и замкнуто, выпуклая функция $f(\mathbf{x}) \in C^{1,1}(Q)$, выполнены (1.2), $\mathbf{x}^* \in Q_* \subset Q \subset E_1^n$. Тогда из равенства*

$$\mathbf{x}^* = P_Q\left[\mathbf{x}^* - \beta\mathbf{B}^{-1}(\mathbf{x}^*)\nabla f(\mathbf{x}^*)\right] \tag{3.3}$$

следует неравенство (3.2).

Д о к а з а т е л ь с т в о. Согласно необходимому и достаточному условию (3.0) проекции $\mathbf{w} = P_Q(\mathbf{v}) \in Q$ в исходной метрике пространства E_1^n при $\mathbf{x}^* \in Q_*$ получаем вариационное неравенство

$$\left(\mathbf{x}^* - (\mathbf{x}^* - \beta\mathbf{B}^{-1}(\mathbf{x}^*)\nabla f(\mathbf{x}^*)), \mathbf{u} - \mathbf{x}^*\right) \geq 0, \ \mathbf{u} \in Q.$$

Отсюда следует неравенство $\beta\,(\mathbf{B}^{-1}(\mathbf{x}^*)\nabla f(\mathbf{x}^*), \mathbf{u}-\mathbf{x}^*) \geq 0,\ \mathbf{u}\in Q$. В силу положительности β, последнее неравенство влечёт (3.2).

Лемма 2 доказана.

Примечание 2. Неравенство (3.2) и равенство (3.3) выражают аналогичный данному в [1] для пространства E^n критерий оптимальности для указанных функций $f(\mathbf{x})$ на выпуклом замкнутом множестве $Q\subset E_1^n$. Следующая лемма выражает связь между необходимыми условиями оптимальности точки $\mathbf{x}^*$ для функции $f(\mathbf{x})$ в исходной метрике пространства E^n ([1], с. 165) и в метрике $\mathbf{B}(\mathbf{x})$ пространства E_1^n в терминах оператора проекции $P_Q^{\mathbf{B}(\mathbf{x})}$.

Лемма 3. *Пусть: 1) множество $Q\subset E_1^n$ выпукло и замкнуто; 2) выпуклая функция $f(\mathbf{x})\in C^{1,1}(Q)$; 3) $Q_*\neq\emptyset$, $\mathbf{x}^*\in Q_*\subset Q\subset E_1^n$.*

Тогда из равенства $\mathbf{x}^* = P_Q^{\mathbf{B}(\mathbf{x}^*)}\left[\mathbf{x}^*-\beta\mathbf{B}^{-1}(\mathbf{x}^*)\nabla f(\mathbf{x}^*)\right]$ следует неравенство

$$(\nabla f(\mathbf{x}^*), \mathbf{u}-\mathbf{x}^*)\geq 0,\ \forall\, \mathbf{u}\in Q. \tag{3.4}$$

Д о к а з а т е л ь с т в о. Из данного равенства, пользуясь в новой метрике пространства E_1^n критерием (2.1) $\mathbf{B}(\mathbf{x})$ – проекции $P_Q^{\mathbf{B}(\mathbf{x})}(\mathbf{v})\in Q$ при $\mathbf{x}=\mathbf{x}^*\in Q_*$ получим неравенство

$$\left(\mathbf{B}(\mathbf{x}^*)\left(\mathbf{x}^*-\mathbf{x}^*+\beta\mathbf{B}^{-1}(\mathbf{x}^*)\nabla f(\mathbf{x}^*)\right), \mathbf{u}-\mathbf{x}^*\right)\geq 0, \quad \mathbf{u}\in Q,$$

которому эквивалентно следующее: $\beta\,(\mathbf{B}(\mathbf{x}^*)\mathbf{B}^{-1}(\mathbf{x}^*)\nabla f(\mathbf{x}^*), \mathbf{u}-\mathbf{x}^*)\geq 0,\ \forall\,\mathbf{u}\in Q$. Отсюда в силу свойств операторов имеем $\beta\,(\nabla f(\mathbf{x}^*), \mathbf{u}-\mathbf{x}^*)\geq 0\ \forall\ \mathbf{u}\in Q$. Это неравенство, поскольку $\beta>0$, влечёт (3.4).

Лемма 3 доказана.

Лемма 4. *В евклидовом пространстве E_1^n имеет место двойное неравенство*

$$\begin{gathered}(1-\varepsilon)\|\mathbf{u}-\mathbf{v}\|^2+(1-\varepsilon^{-1})\|\mathbf{v}-\mathbf{w}\|^2\leq\|\mathbf{u}-\mathbf{w}\|^2\leq\\ \leq(1+\varepsilon)\|\mathbf{u}-\mathbf{v}\|^2+(1+\varepsilon^{-1})\|\mathbf{v}-\mathbf{w}\|^2, \quad \mathbf{u},\mathbf{v},\mathbf{w}\in E_1^n.\end{gathered} \tag{3.5}$$

Д о к а з а т е л ь с т в о. Воспользуемся известным равенством

$$\|\mathbf{u}-\mathbf{w}\|^2=\|\mathbf{u}-\mathbf{v}\|^2+2(\mathbf{u}-\mathbf{v},\mathbf{v}-\mathbf{w})+\|\mathbf{v}-\mathbf{w}\|^2,\ \forall\,\mathbf{u},\mathbf{v},\mathbf{w}\in E_1^n, \tag{3.6}$$

запишем его в форме $\|\mathbf{u}-\mathbf{w}\|^2=\|\mathbf{u}-\mathbf{v}\|^2-2(\mathbf{u}-\mathbf{v},\mathbf{w}-\mathbf{v})+\|\mathbf{v}-\mathbf{w}\|^2$ и второе слагаемое в его правой части оценим с помощью неравенства

$$2|(\mathbf{u}-\mathbf{v},\mathbf{w}-\mathbf{v})|\leq\varepsilon\|\mathbf{u}-\mathbf{v}\|^2+\varepsilon^{-1}\|\mathbf{w}-\mathbf{v}\|^2,\ \varepsilon>0. \tag{3.7}$$

Тогда получим $\|\mathbf{u}-\mathbf{w}\|^2\geq\|\mathbf{u}-\mathbf{v}\|^2-\varepsilon\|\mathbf{u}-\mathbf{v}\|^2-\varepsilon^{-1}\|\mathbf{v}-\mathbf{w}\|^2+\|\mathbf{v}-\mathbf{w}\|^2$. Отсюда следует левая часть неравенства (3.5). Пользуясь (3.7) в равенстве (3.6), получим правую часть (3.5), совпадающую с известным неравенством.

Лемма 4 доказана.

Отметим, что правое неравенство (3.5) известное, здесь оно для единообразия выведено в пространстве E_1^n, а левое неравенство (3.5) является результатом объединения применявшихся ранее при обосновании методов оптимизации разрозненных процедур получения нижней оценки квадрата нормы разности векторов в E^n.

Лемма 5. *Для всякой тройки точек $\mathbf{u},\mathbf{v},\mathbf{x}\in E_1^n$ имеет место неравенство*

$$\|\mathbf{u}-\mathbf{v}\|^2\geq(\varepsilon-1)\|\mathbf{u}-\mathbf{x}\|^2-(1-\varepsilon^{-1})\|\mathbf{v}-\mathbf{x}\|^2, \tag{3.8}$$

где $0 < \varepsilon_1 \le \varepsilon \le \varepsilon_2$, $\varepsilon_{1,2} = \left[s \mp (s^2 - 4l_2 l_3)^{1/2}\right]/(2l_2)$, $l_1 = \|\mathbf{u} - \mathbf{v}\|^2$, $l_2 = \|\mathbf{u} - \mathbf{x}\|^2$, $l_3 = \|\mathbf{v} - \mathbf{x}\|^2$, $s = l_1 + l_2 + l_3$.

Д о к а з а т е л ь с т в о. Поскольку E_1^n метрическое пространство с исходной метрикой $\rho(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| \,\forall\, \mathbf{u}, \mathbf{v} \in E_1^n$, в нём имеет место неравенство ([12], с. 31)

$$|\rho(\mathbf{u}, \mathbf{x}) - \rho(\mathbf{v}, \mathbf{y})| \le \rho(\mathbf{u}, \mathbf{v}) + \rho(\mathbf{x}, \mathbf{y}) \forall\, \mathbf{u}, \mathbf{v}, \mathbf{x}, \mathbf{y} \in E_1^n.$$

Отсюда с помощью формулы для метрики при $\mathbf{x} = \mathbf{y}$ получим $|\|\mathbf{u} - \mathbf{x}\| - \|\mathbf{v} - \mathbf{x}\|| \le \|\mathbf{u} - \mathbf{v}\| + \|\mathbf{x} - \mathbf{x}\|$. Возведём его в квадрат и воспользуемся формулой для квадрата разности, $|\|\mathbf{u} - \mathbf{x}\|^2 - 2\|\mathbf{u} - \mathbf{x}\|\|\mathbf{v} - \mathbf{x}\| + \|\mathbf{v} - \mathbf{x}\|^2| \le \|\mathbf{u} - \mathbf{v}\|^2$, удвоенное произведение под знаком модуля оценим с помощью аналога неравенства (3.7). Тогда придём к неравенству

$$|\|\mathbf{u} - \mathbf{x}\|^2 - \varepsilon\|\mathbf{u} - \mathbf{x}\|^2 - \varepsilon^{-1}\|\mathbf{v} - \mathbf{x}\|^2 + \|\mathbf{v} - \mathbf{x}\|^2| \le \|\mathbf{u} - \mathbf{v}\|^2.$$

Представим абсолютную величину в форме двойного неравенства

$$-\|\mathbf{u} - \mathbf{v}\|^2 \le (1 - \varepsilon)\|\mathbf{u} - \mathbf{x}\|^2 + (1 - \varepsilon^{-1})\|\mathbf{v} - \mathbf{x}\|^2 \le \|\mathbf{u} - \mathbf{v}\|^2. \tag{3.9}$$

Возьмём левую часть этого неравенства и, умножив на -1, придём к неравенству (3.8). Решая неравенство (3.8) относительно $\varepsilon > 0$ при данных обозначениях, получаем для этого числа интервал возможных в (3.8) значений. Заметим, что правое неравенство в (3.9) при $\mathbf{u} = \mathbf{u}$, $\mathbf{x} = \mathbf{v}$, $\mathbf{v} = \mathbf{w}$ совпадает с левым неравенством (3.5).

Лемма 5 доказана.

Примечание 3. Верхняя и нижняя границы числа ε в (3.8) зависят от соотношений длин сторон треугольника с вершинами $\mathbf{v}, \mathbf{u}, \mathbf{x}$ (например, здесь это упорядоченная тройка точек в порядке вычисления по итерационной формуле (2.4) $\mathbf{x}^k, \mathbf{x}^{k+1}, \mathbf{x}^*$; случай их расположения на одной прямой не исключается). Зададим дополнительные ограничения, при которых допустимы конкретные значения ε, используемые в доказательствах теорем. Например, не обременительны условия:

а) $\|\mathbf{u} - \mathbf{x}\| \le \|\mathbf{u} - \mathbf{v}\| \le \|\mathbf{v} - \mathbf{x}\|$, они при $\mathbf{v} = \mathbf{x}^k$, $\mathbf{u} = \mathbf{x}^{k+1}$, $\mathbf{x} = \mathbf{x}^*$ запишутся в виде

$$\|\mathbf{x}^{k+1} - \mathbf{x}^*\| \le \|\mathbf{x}^{k+1} - \mathbf{x}^k\| \le \|\mathbf{x}^k - \mathbf{x}^*\|; \tag{3.10}$$

б) $2\|\mathbf{u} - \mathbf{x}\| \le \|\mathbf{u} - \mathbf{v}\| \le \|\mathbf{v} - \mathbf{x}\|$, их аналог $2\|\mathbf{x}^{k+1} - \mathbf{x}^*\| \le \|\mathbf{x}^{k+1} - \mathbf{x}^k\| \le \|\mathbf{x}^k - \mathbf{x}^*\|$;

в) $(11/5)\|\mathbf{u} - \mathbf{x}\| \le \|\mathbf{u} - \mathbf{v}\| \le \|\mathbf{v} - \mathbf{x}\|$, их аналог $(11/5)\|\mathbf{x}^{k+1} - \mathbf{x}^*\| \le \|\mathbf{x}^{k+1} - \mathbf{x}^k\| \le \|\mathbf{x}^k - \mathbf{x}^*\|$;

В случае а) для вычисления границ ε решаем следующее из (3.8) неравенство

$$l_2\varepsilon^2 - s\varepsilon + l_3 \le 0, \tag{3.11}$$

полагая $l_1 = l_2 = l_3$. Получаем значения $\varepsilon_{1,2} = (3 \mp 5^{1/2})/2$, то есть округлённо можно принять значения из множества $0.39 \le \varepsilon \le 2.61$. В случае б) решаем неравенство, следующее из (3.11), положив $4l_2 = l_1 = l_3$; получаем границы $\varepsilon_{1,2} = (9 \mp 65^{1/2})/2$; тогда приближённый отрезок возможных значений $0.47 \le \varepsilon \le 8.5$. В случае в) решаем неравенство, аналогичное (3.11), положив $(121/25)l_2 = l_1 = l_3$; получаем границы $\varepsilon_{1,2} = (267 \mp 60189^{1/2})/50$; тогда приближённое множество возможных значений $0.47 \le \varepsilon \le 10$, достаточно для применения (3.8) в данной работе.

Отметим, что неравенство (3.8) является новым в обосновании методов решения экстремальных задач и выведено для обоснования методов класса ПОДМ, необходимый математический инструмент при доказательстве сходимости и оценке скорости сходимости методов минимизации. Появление дополнительных ограничений неравенств при его применении здесь представляется естественным. Неравенство (3.8) получено на основе известного из функционального анализа неравенства четырёхугольника для метрики. Его частный случай известен как второе неравенство треугольника для метрики ([14], с. 27) $|\rho(\mathbf{x}, \mathbf{z}) - \rho(\mathbf{y}, \mathbf{z})| \le \rho(\mathbf{x}, \mathbf{y}) \,\forall\, \mathbf{x}, \mathbf{y}, \mathbf{z} \in X$ в метрическом пространстве (X, ρ).

4. Обоснование сходимости метода

Исследуем сходимость ПОДМПМ (2.3), (2.4) с параметрами константами $\alpha_k = \alpha,\ \beta_k = \beta$.

Теорема 1. *Пусть выполнены следующие условия: 1) множество $Q \subset E_1^n$ выпукло и замкнуто; 2) функция $f(\mathbf{x}) \in C^{1,1}(Q)$ выпуклая и выполнены соотношения (1.2); 3) последовательность $\left\{\mathbf{x}^k\right\}$ метода (2.3), (2.4) такова, что имеют место неравенства (3.10в); 4) для оператора $\mathbf{B}(\mathbf{x})\ \forall\ \mathbf{x} \in E_1^n$ с указанными в п.2 свойствами выполняется неравенства (2.5); 5) параметры константы метода семейства (2.3), (2.4) таковы, что*

$$0 < \alpha \le 3/8,\ 0 < \beta \le [4(m^2+m)b - 35(1+\alpha)]/(Lb), \tag{4.1}$$

где $b = 21 - 4\alpha$, $m^2 + m > 5/12$.

Тогда последовательность $\left\{\mathbf{x}^k\right\}$, определяемая методом (2.3), (2.4), (4.1), из любой начальной точки $\mathbf{x}^0 \in E_1^n$ сходится к точке $\mathbf{x}^* \in Q_*$,

$$\|\mathbf{x}^k - \mathbf{x}^*\| \longrightarrow 0, \quad f(\mathbf{x}^k) \longrightarrow f(\mathbf{x}^*), \quad k \to \infty. \tag{4.2}$$

Д о к а з а т е л ь с т в о. Из характеристического свойства (2.1) оператора проектирования в новой метрике пространства E_1^n и (2.4) получим вариационное неравенство

$$\left(\mathbf{B}(\mathbf{z}^k)\left[\mathbf{x}^{k+1} - \mathbf{z}^k + \beta\mathbf{B}(\mathbf{z}^k)^{-1}\nabla f(\mathbf{z}^k)\right], \mathbf{v} - \mathbf{x}^{k+1}\right) \ge 0,\ k \ge 1,\ \mathbf{v} \in Q. \tag{4.3}$$

Это неравенство запишем в форме

$$\begin{gathered}\left(\mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1} - \mathbf{z}^k), \mathbf{x}^{k+1} - \mathbf{z}^k\right) + \left(\mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1} - \mathbf{z}^k), \mathbf{z}^k - \mathbf{v}\right) \le \\ \le \beta\left(\nabla f(\mathbf{z}^k), \mathbf{v} - \mathbf{x}^{k+1}\right),\ k \ge 1,\ \mathbf{v} \in Q.\end{gathered}$$

Полученное из этого при $\mathbf{v} = \mathbf{x}^* \in Q_*$ неравенство сложим с неравенством, полученным из (3.4) при $\mathbf{u} = \mathbf{x}^{k+1}$ и умноженным на $\beta > 0$:

$$\begin{gathered}\left(\mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1} - \mathbf{z}^k), \mathbf{x}^{k+1} - \mathbf{z}^k\right) + \left(\mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1} - \mathbf{z}^k), \mathbf{z}^k - \mathbf{x}^*\right) \le \\ \le \beta\left(\nabla f(\mathbf{z}^k) - \nabla f(\mathbf{x}^*), \mathbf{x}^* - \mathbf{x}^{k+1}\right),\ k \ge 1,\ \mathbf{x}^* \in Q_*.\end{gathered}$$

Здесь в левой части первое слагаемое оценим с помощью (2.5), а в правой части воспользуемся неравенством $(\nabla f(\mathbf{x}) - \nabla f(\mathbf{u}), \mathbf{u} - \mathbf{z}) \le L\|\mathbf{x} - \mathbf{z}\|^2/4,\ \mathbf{x}, \mathbf{u}, \mathbf{z} \in Q$, впервые доказанным в гл.1 работы [2] для $f(\mathbf{x}) \in C^{1,1}(Q)$, положив в нём $\mathbf{x} = \mathbf{z}^k$, $\mathbf{u} = \mathbf{x}^*$, $\mathbf{z} = \mathbf{x}^{k+1}$, то есть $\left(\nabla f(\mathbf{z}^k) - \nabla f(\mathbf{x}^*), \mathbf{x}^* - \mathbf{x}^{k+1}\right) \le L\|\mathbf{x}^{k+1} - \mathbf{z}^k\|^2/4$; тогда

$$\begin{gathered}(m - L\beta/4)\|\mathbf{x}^{k+1} - \mathbf{z}^k\|^2 + \left(\mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1} - \mathbf{z}^k), \mathbf{z}^k - \mathbf{x}^*\right) \le 0, \\ k \ge 1,\ \mathbf{x}^* \in Q_*.\end{gathered} \tag{4.4}$$

В (4.4) сначала оценим второе слагаемое. Рассмотрим тождество

$$\begin{gathered}\|\mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1} - \mathbf{z}^k) - (\mathbf{x}^* - \mathbf{z}^k)\|^2 = \|\mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1} - \mathbf{z}^k)\|^2 - \\ -2\left(\mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1} - \mathbf{z}^k), \mathbf{x}^* - \mathbf{z}^k\right) + \|\mathbf{x}^* - \mathbf{z}^k\|^2.\end{gathered}$$

Обозначив $\mathbf{u} = \mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1} - \mathbf{z}^k)$, $\mathbf{v} = \mathbf{x}^* - \mathbf{z}^k$, запишем

$$2\left(\mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1} - \mathbf{z}^k), \mathbf{z}^k - \mathbf{x}^*\right) = -2(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|^2 - \|\mathbf{u}\|^2 - \|\mathbf{v}\|^2 \tag{4.5}$$

и оценим снизу правую часть; для первого слагаемого с помощью (3.8) при $\mathbf{x}=\mathbf{0}$, $\varepsilon=4$ получим $\|\mathbf{u}-\mathbf{v}\|^2 \ge 3\|\mathbf{u}\|^2-(3/4)\|\mathbf{v}\|^2$; тогда из (4.5), пользуясь (2.5), имеем

$$\begin{gathered}\big(\mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1}-\mathbf{z}^k),\mathbf{z}^k-\mathbf{x}^*\big) \ge 0.5\,(2\|\mathbf{u}\|^2-(7/4)\|\mathbf{v}\|^2) \ge \\ \ge m^2\|\mathbf{x}^{k+1}-\mathbf{z}^k\|^2-(7/8)\|\mathbf{z}^k-\mathbf{x}^*\|^2.\end{gathered}$$

Подставив эту оценку в (4.4), получим

$$a_{01}\|\mathbf{x}^{k+1}-\mathbf{z}^k\|^2 \le (7/8)\|\mathbf{z}^k-\mathbf{x}^*\|^2,\ k\ge 1,\ \mathbf{x}^*\in Q_*, \tag{4.6}$$

где $a_{01}=c-L\beta/4>0$, $\beta<4c/L$, $c=m+m^2$. Здесь проведём преобразования левой части с помощью неравенства (3.8): при $\mathbf{v}=\mathbf{z}^k$, $\mathbf{u}=\mathbf{x}^{k+1}$, $\mathbf{x}=\mathbf{x}^*$, $\varepsilon=5$ имеем

$$0.5\|\mathbf{x}^{k+1}-\mathbf{z}^k\|^2 \ge 2\|\mathbf{x}^{k+1}-\mathbf{x}^*\|^2-(2/5)\|\mathbf{z}^k-\mathbf{x}^*\|^2; \tag{4.7a}$$

при $\mathbf{v}=\mathbf{z}^k$, $\mathbf{u}=\mathbf{x}^{k+1}$, $\mathbf{x}=\mathbf{x}^k$, $\varepsilon=5$ и (2.3) получим

$$\begin{gathered}0.5\|\mathbf{x}^{k+1}-\mathbf{z}^k\|^2 \ge 2\|\mathbf{x}^{k+1}-\mathbf{x}^k\|^2-(2/5)\|\mathbf{z}^k-\mathbf{x}^*\|^2 = \\ = 2\|\mathbf{x}^{k+1}-\mathbf{x}^k\|^2-(2\alpha^2/5)\|\mathbf{y}^k\|^2,\end{gathered}$$

где один квадрат нормы в правой части преобразуем с помощью (3.8) при $\mathbf{v}=\mathbf{x}^k$, $\mathbf{u}=\mathbf{x}^{k+1}$, $\mathbf{x}=\mathbf{x}^*$, $\varepsilon=2$, тогда $\|\mathbf{x}^{k+1}-\mathbf{x}^k\|^2 \ge \|\mathbf{x}^{k+1}-\mathbf{x}^*\|^2-0.5\|\mathbf{x}^k-\mathbf{x}^*\|^2$ и

$$\begin{gathered}0.5\|\mathbf{x}^{k+1}-\mathbf{z}^k\|^2 \ge \|\mathbf{x}^{k+1}-\mathbf{x}^k\|^2+\|\mathbf{x}^{k+1}-\mathbf{x}^*\|^2- \\ -\|\mathbf{x}^k-\mathbf{x}^*\|^2/2-(2\alpha^2/5)\|\mathbf{y}^k\|^2.\end{gathered} \tag{4.7b}$$

Подставив (4.7a) и (4.7b) в (4.6), придём к неравенству

$$\begin{gathered}3a_{01}\|\mathbf{x}^{k+1}-\mathbf{x}^*\|^2+a_{01}\|\mathbf{x}^{k+1}-\mathbf{x}^k\|^2 \le a_{01}\|\mathbf{x}^k-\mathbf{x}^*\|^2/2+ \\ +(2a_{01}\alpha^2/5)\|\mathbf{y}^k\|^2++(7/8+2a_{01}/5)\|\mathbf{z}^k-\mathbf{x}^*\|^2,\ k\ge 1.\end{gathered} \tag{4.8}$$

Правую часть (4.8) преобразуем с помощью неравенства

$$\|\mathbf{z}^k-\mathbf{x}^*\|^2 \le (1+\varepsilon)\|\mathbf{z}^k-\mathbf{x}^k\|^2+(1+\varepsilon^{-1})\|\mathbf{x}^k-\mathbf{x}^*\|^2,\ \varepsilon>0,$$

следующего из правого неравенства (3.5) при $\mathbf{u}=\mathbf{z}^k$, $\mathbf{w}=\mathbf{x}^*$, $\mathbf{v}=\mathbf{x}^k$; положим $\varepsilon=1/\alpha$ и учтём следующее из (2.3) равенство $\|\mathbf{z}^k-\mathbf{x}^k\|^2=\alpha^2\|\mathbf{y}^k\|^2$:

$$a_{02}\|\mathbf{z}^k-\mathbf{x}^*\|^2 \le a_{02}(1+\alpha)\|\mathbf{x}^k-\mathbf{x}^*\|^2+a_{02}(1+\alpha)\alpha\|\mathbf{y}^k\|^2,$$

где $a_{02}=7/8+2a_{01}/5>0$ при $a_{01}>0$. Тогда из (4.8) следует

$$\begin{gathered}3a_{01}\|\mathbf{x}^{k+1}-\mathbf{x}^*\|^2+a_{01}\|\mathbf{x}^{k+1}-\mathbf{x}^k\|^2 \le \\ \le a_{12}\|\mathbf{x}^k-\mathbf{x}^*\|^2+a_{13}\|\mathbf{y}^k\|^2,\ k\ge 1,\end{gathered}$$

где $a_{12}=a_{02}(1+\alpha)+a_{01}/2$, $a_{13}=2a_{01}\alpha^2/5+a_{02}(\alpha+\alpha^2)$. После деления на коэффициент при первом слагаемом имеем

$$\begin{gathered}\|\mathbf{x}^{k+1}-\mathbf{x}^*\|^2+\|\mathbf{x}^{k+1}-\mathbf{x}^k\|^2/3 \le \\ \le a_1\|\mathbf{x}^k-\mathbf{x}^*\|^2+a_2\|\mathbf{y}^k\|^2,\ k\ge 1,\end{gathered} \tag{4.9}$$

где $a_1=a_{12}/(3a_{01})=1/6+a_{02}(1+\alpha)/(3a_{01})$, $a_2=2\alpha^2/15+a_{02}(\alpha+\alpha^2)/(3a_{01})$, $a_1>0$, $a_2>0$ при $a_{01}>0$, $\alpha>0$; $a_1\le 1$ при $0<\beta\le[4cb-35(1+\alpha)]/(Lb)=\beta^{12}$, $b=21-4\alpha$;

$a_2 \le 1/3$ при $\beta \le [8cd - 35(\alpha + \alpha^2)]/(2Ld) = \beta^{13}$, $\beta \le \beta^{12} < \beta^{13}$, $d = 5 - 2\alpha - 4\alpha^2$; $0 < \beta \le \beta^{12} < \beta^{11}$, $0 < \alpha \le 3/8$.

Поскольку $a_1 \le 1$ при условиях теоремы, положим $a_1 = 1$ в (4.9):

$$\begin{aligned} &\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 + \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2/3 \le \\ &\le \|\mathbf{x}^k - \mathbf{x}^*\|^2 + a_2\|\mathbf{y}^k\|^2,\ k \ge 1. \end{aligned} \tag{4.10}$$

Просуммировав неравенства (4.10) от $k = 1$ до $k = m$, получим

$$\begin{gathered} \|\mathbf{x}^{m+1} - \mathbf{x}^*\|^2 + \|\mathbf{x}^{m+1} - \mathbf{x}^m\|^2/3 + (1/3 - a_2)\textstyle\sum_{k=1}^{k=m-1} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 \le \\ \le a_2\|\mathbf{x}^1 - \mathbf{x}^0\|^2 + \|\mathbf{x}^1 - \mathbf{x}^*\|^2. \end{gathered}$$

Отсюда, поскольку $1/3 > 1/3 - a_2 > 0$, $0 < \beta \le \beta^{12} < \beta^{13}$, имеем

$$\begin{gathered} \|\mathbf{x}^{m+1} - \mathbf{x}^*\|^2 + (1/3 - a_2)\textstyle\sum_{k=1}^{k=m} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 \le \\ \le a_2\|\mathbf{x}^1 - \mathbf{x}^0\|^2 + \|\mathbf{x}^1 - \mathbf{x}^*\|^2. \end{gathered}$$

Здесь коэффициенты неотрицательны и

$$\sum_{k=1}^{k=m} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 < \infty\ \forall\, m \ge 1, \quad \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 \to 0,\ k \to \infty,$$

а следовательно, $\|\mathbf{x}^{m+1} - \mathbf{x}^*\|^2 \le a_2\|\mathbf{x}^1 - \mathbf{x}^0\|^2 + \|\mathbf{x}^1 - \mathbf{x}^*\|^2$ и последовательность $\{\mathbf{x}^k\}$ метода (2.3), (2.4), (4.1) ограничена; по теореме Больцано-Вейерштрасса существует точка $\mathbf{v} \in Q$ и подпоследовательность $\{\mathbf{x}^{k_i}\} \to \mathbf{v}$, $i \to \infty$ такая, что

$$\lim_{i\to\infty} \|\mathbf{x}^{k_i} - \mathbf{v}\| = \lim_{i\to\infty} \|\mathbf{x}^{k_i+1} - \mathbf{x}^{k_i}\| = 0. \tag{4.11}$$

Тогда из (2.4) при $k \to \infty$ следует равенство $\mathbf{v} = P_Q^{\mathbf{B}(\mathbf{v})}[\mathbf{v} - \beta[\mathbf{B}(\mathbf{v})]^{-1}\nabla f(\mathbf{v})]$, $\beta > 0$, эквивалентное критерию оптимальности в E_1^n, согласно леммам 1–3. Это значит, что $\mathbf{v} \in Q$ есть точка минимума функции $f(\mathbf{x})$ на множестве Q, то есть $\mathbf{v} = \mathbf{x}^* \in Q_*$.

В силу (4.11) существуют номер $k_0 > 0$ и достаточно малое число $\varepsilon > 0$ такие, что $\forall\, k_i \ge k_0$ выполняются неравенства

$$\|\mathbf{x}^{k_i} - \mathbf{v}\|^2 \le \varepsilon/2,\ \|\mathbf{x}^{k_i+1} - \mathbf{x}^{k_i}\|^2 \le \varepsilon/(2a_2). \tag{4.12}$$

Возьмём числа $N > m \ge k_0$ и просуммируем неравенства (4.10) ещё раз от $k = m$ до $k = N$ при $\mathbf{x}^* = \mathbf{v}$. Получим неравенство

$$\begin{gathered} \|\mathbf{x}^{N+1} - \mathbf{v}\|^2 + \|\mathbf{x}^{N+1} - \mathbf{x}^N\|^2/3 + (1/3 - a_2)\textstyle\sum_{k=m}^{k=N-1} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 \le \\ \le \|\mathbf{x}^m - \mathbf{v}\|^2 + a_2\|\mathbf{x}^m - \mathbf{x}^{m-1}\|^2. \end{gathered}$$

Здесь учтём, что $0 < 1/3 - a_2 < 1/3$ при $0 < \beta \le \beta^{12} < \beta^{13}$, тогда

$$\begin{gathered} \|\mathbf{x}^{N+1} - \mathbf{v}\|^2 + (1/3 - a_2)\textstyle\sum_{k=m}^{k=N} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 \le \\ \le \|\mathbf{x}^m - \mathbf{v}\|^2 + a_2\|\mathbf{x}^m - \mathbf{x}^{m-1}\|^2. \end{gathered}$$

Отсюда при $N \to \infty$ с учётом сходимости ряда и (4.12) имеем:

$$\|\mathbf{x}^{N+1} - \mathbf{v}\|^2 \le \|\mathbf{x}^m - \mathbf{v}\|^2 + a_2\|\mathbf{x}^m - \mathbf{x}^{m-1}\|^2 \le \varepsilon.$$

Это означает фундаментальность последовательности $\{\mathbf{x}^k\}$. Следовательно, $\|\mathbf{x}^k - \mathbf{v}\| \to 0$, $k \to \infty$; не только подпоследовательность, но и вся последовательность $\{\mathbf{x}^k\}$ метода сходится к точке $\mathbf{v} = \mathbf{x}^* \in Q_*$: $\|\mathbf{x}^k - \mathbf{v}\| \to 0$, $k \to \infty$.

Сходимость метода по функционалу покажем с помощью неравенств для функции $f(\mathbf{x}) \in C^{1,1}(Q)$ на выпуклом множестве Q ([1], с. 164) $f(\mathbf{u}) - f(\mathbf{v}) \geq (\nabla f(\mathbf{v}), \mathbf{u} - \mathbf{v})\ \forall\, \mathbf{u}, \mathbf{v} \in Q$ и $(\nabla f(\mathbf{x}^*), \mathbf{u} - \mathbf{x}^*) \geq 0$, $\mathbf{x}^* \in Q_*$, $\mathbf{u} \in Q$ ([1], с. 165). Из неравенств

$$0 \leq (\nabla f(\mathbf{x}^*), \mathbf{u} - \mathbf{x}^*) \leq f(\mathbf{x}^k) - f(\mathbf{x}^*) \leq$$
$$\leq (\nabla f(\mathbf{x}^k), \mathbf{x}^k - \mathbf{x}^*) \leq \|\nabla f_k\| \|\mathbf{x}^k - \mathbf{x}^*\|,$$

доказанной сходимости по аргументу, ограниченности градиента функции, следует второе соотношение из (4.2).

Теорема 1 доказана.

Следствие. *Из теоремы 1 следует монотонность сходимости последовательности* $\{\mathbf{x}^k\}$ *и справедливость неравенств*

$$\|\mathbf{x}^{k-1} - \mathbf{x}^k\| \geq \|\mathbf{x}^k - \mathbf{x}^{k+1}\| \geq \|\mathbf{x}^{k+1} - \mathbf{x}^{k+2}\| \geq \cdots,$$
$$\|\mathbf{x}^{k+1} - \mathbf{x}^*\| \leq \|\mathbf{x}^k - \mathbf{x}^*\| \leq \|\mathbf{x}^{k-1} - \mathbf{x}^*\| \leq \cdots \leq \|\mathbf{x}^0 - \mathbf{x}^*\|.$$

5. Оценки скорости сходимости

5.1. Сначала покажем линейную скорость сходимости метода при условиях теоремы 1 (в теоремах 2 и 3), а затем - сверхлинейную и квадратичную при дополнительных условиях.

Теорема 2. *Пусть выполнены все условия теоремы 1, включая условия для параметров метода. Тогда последовательность* $\{\mathbf{x}^k\}$, *определяемая ПОДМПМ (2.3), (2.4), (4.1), со скоростью геометрической прогрессии сходится к решению задачи (1.1),*

$$\|\mathbf{x}^k - \mathbf{x}^*\| \leq q^k \|\mathbf{x}^0 - \mathbf{x}^*\|, \tag{5.1}$$

где $q = \{[7(1+\alpha)/24 + (0.3 + 2\alpha/15)a_{01}]/a_{01}\}^{1/2}$, $0 < q < 1$ *при условиях (4.1);* $a_{01} = m + m^2 - L\beta/4$.

Доказательство. Заметим, что в условиях данной теоремы все выкладки и рассуждения теоремы 1 и неравенство (4.9) верны.

Покажем, что для слагаемых левой и правой частей (4.9) имеет место неравенство $a_2\|\mathbf{y}^k\|^2 - (1/3)\|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 \leq 0$. Действительно, оно верно в силу неравенств: $\|\mathbf{y}^k\|^2 \leq \|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2$ (в силу следствия теоремы 1); $a_2 \leq 1/3$; $a_2 - 1/3 < a_2$; $(a_2 - 1/3)\left(\|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 - \|\mathbf{y}^k\|^2\right) \geq 0$;

$$a_2\|\mathbf{y}^k\|^2 - (1/3)\|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 \leq (a_2 - 1/3)\left(\|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 - \|\mathbf{y}^k\|^2\right) +$$
$$+ a_2\|\mathbf{y}^k\|^2 - (1/3)\|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 \leq$$
$$\leq a_2\left(\|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 - \|\mathbf{y}^k\|^2\right) + a_2\|\mathbf{y}^k\|^2 - a_2\|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2 \leq 0.$$

Учитывая доказанное неравенство, из (4.9) получаем,

$$\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 \leq q^2\|\mathbf{x}^k - \mathbf{x}^*\|^2,\ k \geq 1, \tag{5.2}$$

где $q = (a_1)^{1/2}$, $0 < q < 1$ при условиях (4.1). Из (5.2) следует (5.1).

Теорема 2 доказана.

Лемма 6. *Пусть последовательность* $\{\mathbf{x}^k\} \to \mathbf{x}^* \in Q_*$ *строится методом класса ПОДМ и выполняются неравенства* $2\|\mathbf{y}^k\| \le \|\mathbf{x}^k - \mathbf{x}^*\| \le \|\mathbf{x}^{k-1} - \mathbf{x}^*\|$ *(следующие из (3.10б)). Тогда для вектора* $\mathbf{y}^k = \mathbf{x}^k - \mathbf{x}^{k-1}$ *приращения аргумента функции* $f(\mathbf{x})$ *имеет место оценка*

$$\|\mathbf{y}^k\| \le (3/\sqrt{8})\|\mathbf{x}^k - \mathbf{x}^*\|. \tag{5.3}$$

Доказательство. Из (3.8) при $\mathbf{u} = \mathbf{x}^k$, $\mathbf{x} = \mathbf{x}^{k+1}$, $\mathbf{v} = \mathbf{x}^*$ следует неравенство $\|\mathbf{x}^k - \mathbf{x}^*\|^2 \ge (\varepsilon - 1)\|\mathbf{y}^k\|^2 - (1 - \varepsilon^{-1})\|\mathbf{x}^{k-1} - \mathbf{x}^*\|^2$. Примем $\varepsilon = 5$ (это допустимо в условиях леммы), тогда $\|\mathbf{x}^k - \mathbf{x}^*\|^2 \ge 4\|\mathbf{y}^k\|^2 - (4/5)\|\mathbf{x}^{k-1} - \mathbf{x}^*\|^2$ или

$$4\|\mathbf{y}^k\|^2 \le \|\mathbf{x}^k - \mathbf{x}^*\|^2 + (4/5)\|\mathbf{x}^{k-1} - \mathbf{x}^*\|^2. \tag{5.4}$$

Из правого неравенства (3.5) при $\mathbf{u} = \mathbf{x}^{k-1}$, $\mathbf{v} = \mathbf{x}^k$, $\mathbf{w} = \mathbf{x}^*$, $\varepsilon = 2/3$ следует $\|\mathbf{x}^{k-1} - \mathbf{x}^*\|^2 \le (5/3)\|\mathbf{y}^k\|^2 + (5/2)\|\mathbf{x}^k - \mathbf{x}^*\|^2$. Подставив его в правую часть (5.4), придём к неравенству $4\|\mathbf{y}^k\|^2 \le \|\mathbf{x}^k - \mathbf{x}^*\|^2 + (4/3)\|\mathbf{y}^k\|^2 + 2\|\mathbf{x}^k - \mathbf{x}^*\|^2$, то есть $8\|\mathbf{y}^k\|^2/3 \le 3\|\mathbf{x}^k - \mathbf{x}^*\|^2$. Отсюда следует неравенство (5.3).

Лемма 6 доказана.

С помощью лемм 3 и 6 можно получить новую оценку линейной скорости сходимости ПОДМПМ (2.3), (2.4) с меньшим, чем полученный в теореме 2, знаменателем прогрессии.

Теорема 3. *Если выполнены все условия теоремы 1 и, кроме того,* $0 < \beta < (16cd - 35e)/(4Ld)$, *то ПОДМПМ (2.3), (2.4)* $\forall\, \mathbf{x}^0 \in E_1^n$ *сходится к точке* $\mathbf{x}^* \in Q_*$ *с оценкой скорости сходимости*

$$\|\mathbf{x}^k - \mathbf{x}^*\| \le q^k\|\mathbf{x}^0 - \mathbf{x}^*\|,\ k \ge 1, \tag{5.5}$$

где $q = (a_1/2 + 9a_2/16 + 1/8)^{1/2}$, $0 < q^2 < 1/2$ *при условиях теоремы,* $d = 27 - 17\alpha - 18\alpha^2$, $e = 8 + 17\alpha + 9\alpha^2$, a_1, a_2, c *из теоремы 1.*

Доказательство. Заметив, что в условиях данной теоремы все выкладки и рассуждения теоремы 1 и неравенство (4.9) верны, в (4.9) второе слагаемое в левой части оценим с помощью (3.8) при $\mathbf{u} = \mathbf{x}^{k+1}$, $\mathbf{v} = \mathbf{x}^k$, $\mathbf{x} = \mathbf{x}^*$, $\varepsilon = 4$:

$$\|\mathbf{x}^{k+1} - \mathbf{x}^k\|^2/3 \ge \|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 - \|\mathbf{x}^k - \mathbf{x}^*\|^2/4.$$

Второе слагаемое в правой части (4.9) оценим с помощью (5.3), $a_2\|\mathbf{y}^k\|^2 \le 9a_2\|\mathbf{x}^k - \mathbf{x}^*\|^2/8$. Подставив эти оценки в (4.9), придём к неравенству $2\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 \le (a_1 + 9a_2/8 + 1/4)\|\mathbf{x}^k - \mathbf{x}^*\|^2$, то есть $\|\mathbf{x}^{k+1} - \mathbf{x}^*\|^2 \le q^2\|\mathbf{x}^k - \mathbf{x}^*\|^2$, $k \ge 1$, где $q^2 = a_1/2 + 9a_2/16 + 1/8$, $0 < q^2 < 1/2$ при $0 < \beta < (16cd - 35e)/(4Ld) < \beta^{12}$, β^{12} из теоремы 1. Отсюда следует (5.5).

Теорема 3 доказана.

5.2. Далее предположим, что функция $f(\mathbf{x}) \in C^{2,1}(E_1^n)$, и покажем, что метод (2.3), (2.4) в этом случае имеет сверхлинейную скорость сходимости. Заметим, что в силу теоремы 1 $\|\mathbf{x}^k - \mathbf{x}^*\| \to 0$, и $\|\mathbf{z}^k - \mathbf{x}^*\| \to 0$ при $\|\mathbf{y}^k\| \to 0$ (следствие теоремы 1), тогда ввиду непрерывности гессиана

$$\|\nabla^2 f(\mathbf{x}^k) - \nabla^2 f(\mathbf{x}^*)\| \to 0,\ \|\nabla^2 f(\mathbf{z}^k) - \nabla^2 f(\mathbf{x}^*)\| \to 0,\ k \to \infty. \tag{5.6}$$

Предположим, что

$$\|\mathbf{B}(\mathbf{z}^k) - \mathbf{B}(\mathbf{x}^*)\| \to 0,\ \|\mathbf{B}(\mathbf{z}^k) - \nabla^2 f(\mathbf{z}^k)\| \to 0,\ k \to \infty. \tag{5.7}$$

Теорема 4. *Пусть выполнены все условия теоремы 1 и, кроме того: 1) функция* $f(\mathbf{x}) \in C^{2,1}(E_1^n)$*; 2) для последовательности* $\{\mathbf{x}^k\} \to \mathbf{x}^* \in Q_*$ *ПОДМПМ (2.3), (2.4)* *существует номер $N>1$ такой, что $\beta_k=1$ при $k\geq N$; 3) выполнены соотношения (5.6)-(5.7). Тогда последовательность $\{\mathbf{x}^k\}$, определяемая методом (2.3), (2.4), (4.1) $\forall\, \mathbf{x}^0\in E_1^n$ со сверхлинейной скоростью сходится к решению задачи (1.1),*

$$\|\mathbf{x}^{k+1}-\mathbf{x}^*\|\leq q_k\|\mathbf{x}^k-\mathbf{x}^*\|,\; q_k\to 0 \text{ при } k\to\infty. \tag{5.8}$$

Д о к а з а т е л ь с т в о. Результаты и выкладки теоремы 1 здесь справедливы. Запишем неравенство (4.3) в форме

$$\begin{gathered}\big(\mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1}-\mathbf{v}),\mathbf{x}^{k+1}-\mathbf{v}\big)\leq\big(\mathbf{B}(\mathbf{z}^k)(\mathbf{z}^k-\mathbf{v}),\mathbf{x}^{k+1}-\mathbf{v}\big)+\\ +\beta\left(\nabla f(\mathbf{z}^k),\mathbf{v}-\mathbf{x}^{k+1}\right),\; k\geq 1,\; \mathbf{v}\in Q.\end{gathered}$$

Полученное из этого при $\mathbf{v}=\mathbf{x}^*\in Q_*$ неравенство сложим с неравенством, полученным из (3.4) при $\mathbf{u}=\mathbf{x}^{k+1}$ и умноженным на $\beta>0$:

$$\begin{gathered}\big(\mathbf{B}(\mathbf{z}^k)(\mathbf{x}^{k+1}-\mathbf{x}^*),\mathbf{x}^{k+1}-\mathbf{x}^*\big)\leq\big(\mathbf{B}(\mathbf{z}^k)(\mathbf{z}^k-\mathbf{x}^*),\mathbf{x}^{k+1}-\mathbf{x}^*\big)+\\ +\beta\left(\nabla f(\mathbf{z}^k)-\nabla f(\mathbf{x}^*),\mathbf{x}^*-\mathbf{x}^{k+1}\right),\; k\geq 1,\; \mathbf{x}^*\in Q_*.\end{gathered}$$

Здесь левую часть оценим снизу с помощью (2.5), а в правой части во втором скалярном произведении воспользуемся формулой Лагранжа,

$$\begin{gathered}m\|\mathbf{x}^{k+1}-\mathbf{x}^*\|^2\leq\big(\mathbf{B}(\mathbf{z}^k)(\mathbf{z}^k-\mathbf{x}^*),\mathbf{x}^{k+1}-\mathbf{x}^*\big)+\\ +\beta\left(\nabla^2 f(\xi^k)(\mathbf{z}^k-\mathbf{x}^*),\mathbf{x}^*-\mathbf{x}^{k+1}\right),\\ k\geq 0,\; \xi^k=\mathbf{z}^k-\theta(\mathbf{z}^k-\mathbf{x}^*),\theta\in[0;1].\end{gathered}\tag{5.9}$$

Преобразуем (5.9), $m\|\mathbf{x}^{k+1}-\mathbf{x}^*\|^2\leq\big(\mathbf{B}(\mathbf{z}^k)-\beta\nabla^2 f(\xi^k)(\mathbf{z}^k-\mathbf{x}^*),\mathbf{x}^{k+1}-\mathbf{x}^*\big)$, и воспользуемся неравенством Коши-Буняковского, тогда придём к неравенству

$$m\|\mathbf{x}^{k+1}-\mathbf{x}^*\|\leq\|\mathbf{B}(\mathbf{z}^k)-\beta\nabla^2 f(\xi^k)\|\|\mathbf{z}^k-\mathbf{x}^*\|. \tag{5.10}$$

Последний сомножитель в правой части (5.10) оценим с помощью (2.3) и (5.3),

$$\begin{gathered}\|\mathbf{z}^k-\mathbf{x}^*\|=\|\mathbf{x}^k+\alpha\mathbf{y}^k-\mathbf{x}^*\|\leq\\ \leq\|\mathbf{x}^k-\mathbf{x}^*\|+\alpha\|\mathbf{y}^k\|\leq(1+3\alpha/\sqrt{8})\|\mathbf{x}^k-\mathbf{x}^*\|.\end{gathered}\tag{5.11}$$

Подставим оценку (5.11) в (5.10) и учтём, что $\beta_k=1$ при $k\geq N$; тогда получим

$$m\|\mathbf{x}^{k+1}-\mathbf{x}^*\|\leq\|\mathbf{B}(\mathbf{z}^k)-\nabla^2 f(\xi^k)\|(1+3\alpha/\sqrt{8})\|\mathbf{x}^k-\mathbf{x}^*\|. \tag{5.12}$$

Из (5.12) следует оценка

$$\|\mathbf{x}^{k+1}-\mathbf{x}^*\|\leq q_k\|\mathbf{x}^k-\mathbf{x}^*\|,\; q_k\to 0 \text{ при } k\to\infty \tag{5.13}$$

где $q_k=\|\mathbf{B}(\mathbf{z}^k)-\nabla^2 f(\xi^k)\|(1+3\alpha/\sqrt{8})/m$,

$$\|\mathbf{B}(\mathbf{z}^k)-\nabla^2 f(\xi^k)\|\leq\|\mathbf{B}(\mathbf{z}^k)-\nabla^2 f(\mathbf{x}^*)\|\to 0 \text{ при } k\to\infty. \tag{5.14}$$

Из (5.13), (5.14) следует (5.8).

Теорема 4 доказана.

5.3. При дополнительном условии относительно операторов $\mathbf{B}_k$ получим оценку квадратичной сходимости ПОДМПМ (2.3), (2.4). Воспользуемся обобщением неравенства из работы [11]. Предположим, что константа $c_1 > 0$ такова, что $\forall\, k \geq N$ имеет место неравенство

$$\|\mathbf{B}(\mathbf{z}^k) - \nabla^2 f(\mathbf{x}^*)\| \leq c_1 \|\mathbf{z}^k - \mathbf{x}^*\|. \qquad (5.15)$$

Теорема 5. *Пусть выполнены все условия теорем 1 и 4 и (5.15). Тогда последовательность* $\{\mathbf{x}^k\}$ *метода (2.3), (2.4), (4.1)* $\forall\, \mathbf{x}^0 \in E_1^n$ *с квадратичной скоростью сходится к решению* $\mathbf{x}^* \in Q_*$ *задачи (1.1),*

$$\|\mathbf{x}^{k+1} - \mathbf{x}^*\| \leq c\|\mathbf{x}^k - \mathbf{x}^*\|^2,\ c = c_1(1 + 3\alpha/\sqrt{8})^2/m. \qquad (5.16)$$

Доказательство. Заметим, что при условиях теоремы 5 все выкладки теорем 1 и 4, а также неравенство (5.12), справедливы. Воспользуемся в (5.12) неравенствами (5.6), (5.7), (5.14), (5.15), а также оценками (5.11) и

$$\|\mathbf{B}(\mathbf{z}^k) - \nabla^2 f(\xi^k)\| \leq \|\mathbf{B}(\mathbf{z}^k) - \nabla^2 f(\mathbf{x}^*)\| \leq$$
$$\leq c_1\|\mathbf{z}^k - \mathbf{x}^*\| \leq c_1(1 + 3\alpha/\sqrt{8})\|\mathbf{x}^k - \mathbf{x}^*\|.$$

Тогда из (5.12) получим $\|\mathbf{x}^{k+1} - \mathbf{x}^*\| \leq c_1(1 + 3\alpha/\sqrt{8})^2\|\mathbf{x}^k - \mathbf{x}^*\|^2/m$, $k \geq N$. Отсюда следует (5.16).

Теорема 5 доказана.

Примечание 5. Для численных реализаций представляет интерес также и модификация ПОДМПМ (2.3), (2.4):

$$1 \text{ этап.} \mathbf{y}^k = \mathbf{x}^k - \mathbf{x}^{k-1},\ \mathbf{z}^k = \mathbf{x}^k + \alpha_k \mathbf{y}^k/\|\mathbf{y}^k\|, \mathbf{x}^0, \mathbf{x}^1 \in E_1^n,\ f(\mathbf{x}^0) > f(\mathbf{x}^1);$$

$$2 \text{ этап.}\quad \mathbf{x}^{k+1} = P_Q^{\mathbf{B}(\mathbf{z}^k)}\left(\mathbf{z}^k + \beta_k \mathbf{p}(\mathbf{z}^k)/\|\nabla f(\mathbf{z}^k)\|\right),\ k = 1, 2, ...,$$

где $\mathbf{p}(\mathbf{z}^k) = -\mathbf{B}(\mathbf{z}^k)^{-1}\nabla f(\mathbf{z}^k)$ (см. п. 2), параметры метода α_k, β_k константы или вычисляются на итерациях по заданным правилам. Например, α_k можно выбирать из условия, что точка $\mathbf{z}^k$ лежит внутри односвязной замкнутой области, ограниченной гиперповерхностью уровня $f(\mathbf{x}) = const = f(\mathbf{x}^{k-1})$, а β_k вычисляется одномерной минимизацией, $\beta_k = argmin_{\beta>0} f\left(\mathbf{z}^k - \beta\mathbf{p}(\mathbf{z}^k)/\|\nabla f(\mathbf{z}^k)\|\right)$; другие обозначения те же, что и в (2.3), (2.4). Доказательство сходимости и оценки скорости сходимости метода аналогичны данным для ПОДМПМ (2.3), (2.4).

Список литературы

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

2. Антипин А. С., “Непрерывные и итеративные процессы с операторами проектирования и типа проектирования”, *Вопросы кибернетики. Вычислительные вопросы анализа больших систем*, 1989, 5–43..

3. Амочкина Т. В., Недич А., “Об одном варианте непрерывного метода проекции градиента второго порядка и его дискретном аналоге”, *Вестник МГУ, Сер. 15. Вычисл. матем. и кибернет.*, 1995, № 2, 5–11..

4. Малинов В. Г., “Четырехпараметрические двухшаговые проекционные методы минимизации первого порядка”, *Журнал вычислительной математики и математической физики*, **36**:12 (1996), 48–56..

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

6. Амочкина Т. В., "Непрерывный метод проекции градиента второго порядка с переменной метрикой", *Журнал вычислительной математики и математической физики*, **37**:10 (1997), 1174–1182..

7. Малинов В. Г., "О проекционном квазиньютоновском обобщенном двухшаговом методе минимизации и оптимизации траектории летательного аппарата", *Журнал Средневолжского математического общества*, **12**:4 (2010), 37–48..

8. Malinov V. G., "Projection Two-step Variable Metric Methods", *4th Moscow International Conference on Operations Research (ORM2004). Moscow. September 21-24, 2004. Proceedings*, 2004, 135–137..

9. Wolfe M. A., "A Quasi-Newton Method with Memory for Unconstrained Function Minimization", *J. Inst. Mathematics & Applications*, **15**:1 (1975), 85–94..

10. Ford J. A., Moghrabi I. A., "Multi-step quasi-Newton methods for optimization", *Journal of Computational and Applied Mathematics*, **50**:Complete (1994), 305–323..

11. Conn A. R., Gould N. I. M., Toint Ph. L., "Convergence of quasi-Newton matrices generated by the symmetric rank one update", *Mathematical Programming*, **50**:2 (1991), 177–195..

12. Канторович Л. В., Акилов Г. П., *Функциональный анализ*, Наука, М., 1977, 744. с.

13. Малинов В.Г., "Проекционный двухшаговый МПМ и численное решение задачи оптимального управления", *Журнал Средневолжского математического общества*, **14**:1 (2012), 83–91..

14. Шилов Г. Е., *Математический анализ. Специальный курс*, Физматлит, М., 1960, 388. с.

PGTM with projecting in variable metric

© **V. G. Malinov** [2]

Abstract. In the work projection generalized two-step two-stage method (PGTM) with projecting in variable metric (PGTVMM) for solving finite dimensional minimization problems on the convex closed set in the Euclidean space is proposed. It may be used as well for solution of functional equations and other problems. The convergence of the method is proved for continuously differentiable convex functions with a Lipschitz gradients. Estimates rate of convergence are proved: at first linear rate of convergence for convex smooth functions, afterwards derived superlinear and quadratic rate of convergence for twice differentiable functions on supplemental suppositions.

Key Words: PGTVMM, projecting in variable metric, convergence, rate of convergence.

[2] Assistant Professor of Ulyanovsk State University, Ulyanovsk; vgmalinov@mail.ru.