Том 30, № 152 2025

НАУЧНАЯ СТАТЬЯ

https://doi.org/10.20310/2686-9667-2025-30-152-346-360

УДК 519.6

Глобализованный кусочный метод Левенберга–Марквардта с процедурой для предотвращения сходимости к нестационарным точкам

Алексей Феридович ИЗМАИЛОВ, Чжибай ЯНЬ

ФГБОУ ВО «Московский государственный университет им. М. В. Ломоносова»

119991, ГСП-2, Российская Федерация, г. Москва, Ленинские горы, 1

Аннотация. Современные версии метода Левенберга–Марквардта для уравнений с ограничениями обладают сильными свойствами локальной сверхлинейной сходимости, допускающими возможную неизолированность решений и возможную негладкость уравнений. Недавно был разработан соответствующий глобальной сходящийся вариант алгоритма для кусочно-гладкого случая, основанный на одномерном поиске для квадрата невязки в евклидовой норме. Для этого алгоритма была показана глобальная сходимость к стационарным точкам для какого-то активного гладкого кусочного отображения, причем примеры показывают, что установить более сильные свойства глобальной сходимости для этого алгоритма без дальнейших его модификаций невозможно. В этой статье разрабатывается такая модификация глобализованного кусочного метода Левенберга–Марквардта, позволяющая избегать нежелательных предельных точек, тем самым обеспечивая желаемое свойство B-стационарности предельных точек для задачи минимизации квадрата невязки исходного уравнения в евклидовой норме, на множестве, задаваемом ограничениями. Конструкция состоит в идентификации гладких кусочных отображений, активных в потенциальных предельных точках, посредством использования подходящей оценки расстояния для активного гладкого кусочного отображения, используемого на текущей итерации, с последующим переключением, при необходимости, на более перспективное идентифицированное кусочное отображение. Устанавливаются глобальная сходимость к B-стационарным точкам и асимптотическая сверхлинейная скорость сходимости, где последнее также основано на подходящей оценке расстояния, но в этом случае до решений исходного уравнения с ограничениями.

Ключевые слова: кусочно-гладкое уравнение, уравнение с ограничением, кусочный метод Левенберга–Марквардта, глобальная сходимость, сверхлинейная сходимость

Благодарности: Исследование выполнено за счет гранта Российского научного фонда (проект № 24-21-00015, https://rscf.ru/project/24-21-00015/).

Для цитирования: *Измаилов А.Ф., Янь Ч.* Глобализованный кусочный метод Левенберга–Марквардта с процедурой для предотвращения сходимости к нестационарным точкам // Вестник российских университетов. Математика. 2025. Т. 30. № 152. С. 346–360. https://doi.org/10.20310/2686-9667-2025-30-152-346-360

SCIENTIFIC ARTICLE

https://doi.org/10.20310/2686-9667-2025-30-152-346-360

Globalized piecewise Levenberg–Marquardt method with a procedure for avoiding convergence to nonstationary points

Alexey F. IZMAILOV, Zhibai YAN
Lomonosov Moscow State University
1 Leninskiye Gory, Moscow 119991, Russian Federation

Abstract. The modern version of the Levenberg–Marquardt method for constrained equations possess strong properties of local superlinear convergence, allowing for possibly nonisolated solutions and possibly nonsmooth equations. A related globally convergent variant of the algorithm for the piecewise-smooth case, based on linesearch for the squared Euclidian norm residual, has recently been developed. Global convergence of this algorithm to stationary points for some active smooth selections has been shown, and examples demonstrate that no any stronger global convergence properties can be established for this algorithm without further modifications. In this paper, we develop such a modification of the globalized piecewise Levenberg–Marquardt method, that avoids undesirable accumulation points, thus achieving the intended property of B-stationarity of accumulation points for the problem of minimization of the squared Euclidian norm residual of the original equation over the constraint set. The construction consists of identifying smooth selections active at potential accumulation points by means of an appropriate error bound for an active smooth selection employed at the current iteration, and then switching to a more promising identified selection when needed. Global convergence to B-stationary points and asymptotic superlinear convergence rate are established, the latter again relying on an appropriate error bound property, but this time for the solutions of the original constrained equation.

Keywords: piecewise smooth equation, constrained equation, piecewise Levenberg–Marquardt method, global convergence, superlinear convergence

Acknowledgements: The research was supported by the Russian Science Foundation (project no. 24-21-00015, https://rscf.ru/en/project/24-21-00015/).

Mathematics Subject Classification: 47J05, 49M15, 65H10, 90C33.

For citation: Izmailov A.F., Yan Z. Globalized piecewise Levenberg–Marquardt method with a procedure for avoiding convergence to nonstationary points. *Vestnik rossiyskikh universitetov. Matematika = Russian Universities Reports. Mathematics*, **30**:152 (2025), 346–360.
https://doi.org/10.20310/2686-9667-2025-30-152-346-360

Введение

Метод Левенберга–Марквардта для уравнений с ограничениями в виде включения в выпуклое замкнутое множество является признанным средством решения таких задач, обладающим сильными свойствами локальной сверхлинейной сходимости, в том числе в случае неизолированных решений, а также в случаях возможной негладкости уравнения; см. недавний обзор в [1] и цитированную в нем литературу. Для кусочно-гладких уравнений с ограничениями в [2] был предложен кусочный метод Левенберга–Марквардта, а в [3] был разработан соответствующий глобализованный алгоритм и обоснованы свойства его глобальной и асимптотической сверхлинейной сходимости. В частности, была показана глобальная сходимость данного алгоритма к стационарным точкам для какого-то активного гладкого кусочного отображения. При этом для алгоритмов такого рода идеальным был бы результат о B-стационарности предельных точек для задачи минимизации квадрата невязки исходного уравнения в евклидовой норме, на множестве, задаваемом ограничениями. Однако, без дальнейших модификаций алгоритма такое свойство гарантировать нельзя; см. примеры в [3] и разд. 4. ниже.

В этой статье разрабатывается модификация глобализованного кусочного метода Левенберга–Марквардта, позволяющая избегать нежелательных предельных точек, тем самым обеспечивая желаемое свойство B-стационарности. Процедура состоит в идентификации гладких кусочных отображений, активных в потенциальных предельных точках, посредством использования подходящей оценки расстояния для активного гладкого кусочного отображения, используемого на текущей итерации, с последующим переключением, при необходимости, на более перспективное идентифицированное кусочное отображение. Устанавливаются глобальная сходимость к B-стационарным точкам и асимптотическая сверхлинейная скорость сходимости, где последнее также основано на подходящей оценке расстояния, но в этом случае до решений исходного уравнения с ограничениями.

Статья организована следующим образом. В разд. 1. приводится постановка задачи, а также излагается кусочный метод Левенберга–Марквардта для ее решения, а также алгоритм, реализующий глобализацию сходимости этого метода. Обсуждаются имеющиеся результаты о локальной сходимости и скорости сходимости, а также о глобальной сходимости данного алгоритма. Разд. 2. посвящен идентификации гладких кусочных отображений активных в потенциальной предельной точке генерируемой алгоритмом последовательности, и характеризациям используемой для этого оценки расстояния до стационарной точки, соответствующей текущему кусочному отображению. В разд. 3. излагается модификация глобализованного алгоритма, снабженная процедурой, позволяющей избегать нежелательных предельных точек, а также теория глобальной сходимости и асимптотической сверхлинейной сходимости модифицированного алгоритма. Наконец, в разд. 4. приводятся примеры, демонстрирующие возникновение решаемой в этой статье проблемы, а также ее преодоление посредством предложенной процедуры.

Используемые ниже обозначения вполне традиционны. Через $\ker A$ обозначается ядро (множество нулей) линейного оператора A. Всюду $\|\cdot\|$ — это евклидова норма, а $\|\cdot\|_\infty$ — норма, определяемая как максимум модулей компонент. Расстояние от точки $u \in \mathbb{R}^p$ до множества $U \subset \mathbb{R}^p$ определяется как $\operatorname{dist}(u, U) = \inf_{\widetilde{u} \in U} \|u - \widetilde{u}\|$, а единственная проекция u на замкнутое выпуклое множество $P \subset \mathbb{R}^p$ обозначается $\pi_P(u)$. Для последовательности $\{u^k\} \subset \mathbb{R}^p$, сходящейся к некоторому u, скорость сходимости называется сверхлинейной с Q-порядком $\tau > 0$, если существует $c > 0$ такое, что

$$\|u^{k+1} - u\| \leqslant c\|u^k - u\|^\tau$$

для всех достаточно больших k.

1. Кусочно-гладкие уравнения с ограничениями и кусочный метод Левенберга–Марквардта

Данная статья посвящена численным методам решения уравнения с ограничением

$$\Phi(u) = 0, \quad u \in P, \tag{1.1}$$

где $\Phi : \mathbb{R}^p \to \mathbb{R}^q$ — заданное отображение, а $P \subset \mathbb{R}^p$ — заданное множество. В этой работе всюду предполагается, что P выпукло и замкнуто, и что отображение Φ является кусочно-гладким, т. е. оно непрерывно, и существует конечный набор гладких (непрерывно дифференцируемых) кусочных отображений $\Phi^1, \ldots, \Phi^s : \mathbb{R}^p \to \mathbb{R}^q$, такой, что

$$\Phi(u) \in \{\Phi^1(u), \ldots, \Phi^s(u)\} \quad \forall\, u \in \mathbb{R}^p.$$

Для каждого $u \in \mathbb{R}^p$ определим множество

$$\mathcal{A}(u) = \{j \in \{1, \ldots, s\} \mid \Phi(u) = \Phi^j(u)\} \tag{1.2}$$

индексов гладких кусочных отображений активных в точке u. Далее будем считать фиксированным произвольное отображение $J : \mathbb{R}^p \to \mathbb{R}^{q\times p}$, удовлетворяющее условию

$$J(u) \in \{(\Phi^j)'(u) \mid j \in \mathcal{A}(u)\} \quad \forall\, u \in \mathbb{R}^p. \tag{1.3}$$

Для текущего приближения $u^k \in P$ введенный в [2] кусочный метод Левенберга–Марквардта (LM) с ограничениями определяет следующее приближение как $u^k + v^k$, где смещение v^k есть решение подзадачи

$$\frac{1}{2}\|\Phi(u^k) + J(u^k)v\|^2 + \frac{1}{2}\sigma(u^k)\|v\|^2 \to \min, \quad u^k + v \in P, \tag{1.4}$$

с некоторой функцией $\sigma : P \to \mathbb{R}_+$, определяющей значения параметра регуляризации.

Локальная сверхлинейная сходимость метода обоснована в [2, теорема 2.1]. Помимо того, что производные гладких кусочных отображений $\Phi^1, \ldots, \Phi^s$ удовлетворяют условию Липшица вблизи решения $\bar{u}$ задачи (1.1), этот результат использует следующие два предположения:

– P-свойство, введенное в [4, с. 434]: существует окрестность U точки $\bar{u}$ такая, что

$$\forall\, j \in \mathcal{A}(\bar{u}) \quad (\Phi^j)^{-1}(0) \cap P \cap U \subset \Phi^{-1}(0) \cap P; \tag{1.5}$$

– кусочная оценка расстояния

$$\forall\, j \in \mathcal{A}(\bar{u}) \quad \operatorname{dist}(u, (\Phi^j)^{-1}(0) \cap P) = O(\|\Phi^j(u)\|) \tag{1.6}$$

при $u \in P$ стремящемся к $\bar{u}$.

Тогда для любого $\delta > 0$ и любого $u^0 \in P$, достаточно близкого к $\bar{u}$, описанный выше метод LM, использующий

$$\sigma(u^k) = \|\Phi(u^k)\|^\theta \tag{1.7}$$

с некоторым фиксированным $\theta \in (0, 2]$, однозначно определяет последовательность $\{u^k\} \subset B(\bar{u}, \delta)$, и эта последовательность сходится к некоторому решению задачи (1.1), причем скорость сходимости сверхлинейная с Q-порядком $\min\{\theta + 1, 2\}$ (либо u^k совпадает с таким решением после конечного числа шагов).

Отметим, что предположения указанного результата допускают как негладкость отображения Φ, так и неизолированность решения $\bar{u}$ задачи (1.1).

Глобализация сходимости метода LM реализована в следующем алгоритме, предложенном в [3, алгоритм 3.2]. Введем функцию $\varphi : \mathbb{R}^p \to \mathbb{R}$,

$$\varphi(u) = \frac{1}{2}\|\Phi(u)\|^2.$$

Алгоритм 1. Фиксируем отображение $J : \mathbb{R}^p \to \mathbb{R}^{q\times p}$, удовлетворяющее (1.3). Фиксируем параметры $\theta > 0$, $\varepsilon \in (0, 1)$ и $\varkappa \in (0, 1)$. Выбираем $u^0 \in P$ и полагаем $k = 0$.

1. Если $\Phi(u^k) = 0$, стоп. Иначе вычисляем $\sigma(u^k)$ согласно (1.7).

2. Вычисляем v^k как решение подзадачи (1.4). Если $v^k = 0$, стоп.

3. Полагаем $\alpha = 1$. Если выполняется неравенство

$$\varphi(u^k + \alpha v^k) \leqslant \varphi(u^k) - \varepsilon\alpha\sigma(u^k)\|v^k\|^2, \tag{1.8}$$

полагаем $\alpha_k = \alpha$. Иначе заменяем α на $\kappa\alpha$ до тех пор, пока неравенство (1.8) не будет выполнено.

4. Полагаем $u^{k+1} = u^k + \alpha_k v^k$, увеличиваем k на 1, и переходим к шагу 1.

Свойства глобальной сходимости данного алгоритма установлены в [3, теорема 3.3] (см. также [3, замечание 3.1]). А именно, при выполнении условия

$$\|\Phi(u)\| \leqslant \|\Phi^j(u)\| \quad \forall\, j \in \{1, \ldots, s\},\ \forall\, u \in P, \tag{1.9}$$

алгоритм однозначно определяет последовательность $\{u^k\}$, любая предельная точка $\bar{u}$ которой удовлетворяет, *для некоторого* $j \in \mathcal{A}(\bar{u})$, условию стационарности

$$\langle \varphi_j'(\bar{u}),\, u - \bar{u}\rangle \geqslant 0 \quad \forall\, u \in P \tag{1.10}$$

в задаче оптимизации

$$\varphi_j(u) \to \min, \quad u \in P, \tag{1.11}$$

где для всякого $j \in \{1, \ldots, s\}$ гладкая функция $\varphi_j : \mathbb{R}^p \to \mathbb{R}$ определяется равенством

$$\varphi_j(u) = \frac{1}{2}\|\Phi^j(u)\|^2.$$

Здесь имеется в виду, что если алгоритм 1 останавливается (на шаге 1 или 2) с текущим приближением u^k после конечного числа итераций, то $\bar{u} = u^k$.

Приведенный результат о глобальной сходимости не идеален: идеальный результат для методов такого рода должен был бы состоять в B-стационарности предельной точки $\bar{u}$, что означает выполнение

$$\varphi'(\bar{u};\, u - \bar{u}) \geqslant 0 \quad \forall\, u \in P,$$

где через $\varphi'(\bar{u};\, v)$ обозначается производная функции φ в точке $\bar{u}$ по направлению $v \in \mathbb{R}^p$. Однако, если $\mathcal{A}(\bar{u})$ состоит более чем из одного элемента, то B-стационарность может не иметь места при выполнении (1.10) для какого-то $j \in \mathcal{A}(\bar{u})$, и более того, такие точки $\bar{u}$ могут притягивать последовательности, генерируемые алгоритмом 1, из широких областей начальных точек; см. [3, пример 3.1]. Согласно [3, предложение 3.2], при выполнении условия (1.9), B-стационарность равносильна выполнению (1.10) *для всех* $j \in \mathcal{A}(\bar{u})$. Соответственно, целью дальнейшего является модификация алгоритма 1, обеспечивающая выполнение последнего свойства в предельных точках генерируемых последовательностей.

2. Оценка расстояния и идентификация активных кусочных отображений

Пусть $\widehat{\jmath} \in \{1,\, \ldots,\, s\}$ — фиксированный индекс гладкого кусочного отображения $\Phi^{\widehat{\jmath}}$. (Имеется в виду, что для текущего приближения u^k это тот индекс $\widehat{\jmath} \in \mathcal{A}(u^k)$, для которого $J(u^k) = (\Phi^{\widehat{\jmath}})'(u^k)$; см. (1.3).) Конструкция процедуры, рассматриваемой в следующем разделе, основана на асимптотической идентификации гладких кусочных отображений, являющихся активными в потенциальных предельных точках последовательности $\{u^k\}$, по информации, доступной в текущем приближении u^k. Такая идентификация использует оценки расстояния до потенциальных предельных точек.

В [5] в контексте метода LP-Newton оценивается расстояние до C-стационарных (или, что в данном случае то же самое [6, предложение 3.2], B-стационарных) точек задачи оптимизации

$$f_{\widehat{\jmath}}(u) \to \min, \quad u \in P,$$

где функции $f_j : \mathbb{R}^p \to \mathbb{R}$, $j \in \{1,\, \ldots,\, s\}$, определяются равенством

$$f_j(u) = \|\Phi^j(u)\|_\infty.$$

Эти функции негладкие из-за негладкости ∞-нормы, и в [5, теорема 3.1] для оценки расстояния используется функция $\Delta(\cdot)$, значения которой вычисляются на итерациях глобализованного метода LP-Newton.

В контексте метода LM нужно оценивать расстояние до стационарных точек задачи оптимизации (1.11) с $j = \widehat{\jmath}$. Естественная мера нестационарности точки $u \in P$ при этом имеет вид $r_{\widehat{\jmath}}(u)$, где $r_j : \mathbb{R}^p \to \mathbb{R}_+$,

$$r_j(u) = \|u - \pi_P(u - \varphi_j'(u))\|, \tag{2.1}$$

$$\varphi_j'(u) = ((\Phi^j)'(u))^\top \Phi^j(u). \tag{2.2}$$

Требуемая оценка расстояния вблизи стационарной точки $\bar{u}$ задачи (1.11) с $j = \widehat{\jmath}$ состоит в выполнении

$$u - \bar{u} = O(r_{\widehat{\jmath}}(u)) \tag{2.3}$$

при $u \in P$ стремящемся к $\bar{u}$. Напомним, что стационарность точки $\bar{u}$ в такой задаче характеризуется равенством $r_{\widehat{\jmath}}(u) = 0$, что равносильно выполнению условий $\bar{u} \in P$ и (1.10) с $j = \widehat{\jmath}$.

Согласно [7, предложение 1.31], оценка расстояния (2.3) при $u \to \bar{u}$ (без требования $u \in P$) равносильна так называемой полуустойчивости $\bar{u}$ как решения вариационного неравенства

$$u \in P, \quad \langle \varphi'_{\widehat{\jmath}}(u),\, \widetilde{u} - u \rangle \geqslant 0 \quad \forall\, \widetilde{u} \in P, \tag{2.4}$$

что, согласно [7, предложение 1.33], в случае двукратной дифференцируемости $\varphi_{\widehat{\jmath}}$ в точке $\bar{u}$ и полиэдральности множества P имеет место тогда и только тогда, когда решение $v = 0$ аффинного вариационного неравенства

$$\bar{u} + v \in P, \quad \langle \varphi'_{\widehat{\jmath}}(\bar{u}) + \varphi''_{\widehat{\jmath}}(\bar{u})v,\, u - \bar{u} - v \rangle \geqslant 0 \quad \forall\, u \in P,$$

изолировано.

Дальнейшие расшифровки приведенной характеризации оценки расстояния (2.3) возможны для более специальных множеств P, например, когда P является параллелепипедом, т. е. имеет вид

$$P = \{u \in \mathbb{R}^p \mid a_i \leqslant u_i \leqslant b_i,\, i \in \{1,\, \ldots,\, p\}\} \tag{2.5}$$

с некоторыми a_i и b_i такими, что $-\infty \leqslant a_i < b_i \leqslant +\infty$, $i \in \{1,\, \ldots,\, p\}$. В частности, если $P = \mathbb{R}^p_+$, то из (2.1) следует, что

$$r_{\widehat{\jmath}}(u) = \|\min\{u,\, \varphi'_{\widehat{\jmath}}(u)\}\|,$$

где минимум берется покомпонентно, а стационарность точки $\bar{u}$ в задаче (1.11) означает, что $\bar{u}$ является решением нелинейной комплементарной задачи

$$u \geqslant 0, \quad \varphi'_{\widehat{\jmath}}(u) \geqslant 0, \quad \langle u,\, \varphi'_{\widehat{\jmath}}(u) \rangle = 0.$$

Согласно [7, предложение 1.34], полуустойчивость $\bar{u}$, а значит, и оценка расстояния (2.3) при этом имеют место тогда и только тогда, когда система

$$\begin{gathered} v_i \geqslant 0, \quad (\varphi''_{\widehat{\jmath}}(\bar{u})v)_i \geqslant 0, \quad v_i(\varphi''_{\widehat{\jmath}}(\bar{u})v)_i = 0, \quad i \in \{1,\, \ldots,\, p\} : \, \bar{u}_i = 0 = (\varphi'_{\widehat{\jmath}}(\bar{u}))_i, \\ (\varphi''_{\widehat{\jmath}}(\bar{u})v)_i = 0, \quad i \in \{1,\, \ldots,\, p\} : \, \bar{u}_i > 0 = (\varphi'_{\widehat{\jmath}}(\bar{u}))_i, \\ v_i = 0, \quad i \in \{1,\, \ldots,\, p\} : \, \bar{u}_i = 0 < (\varphi'_{\widehat{\jmath}}(\bar{u}))_i, \end{gathered}$$

имеет единственное решение $v = 0$.

Если же $P = \mathbb{R}^p$, то из (2.1) следует, что

$$r_{\widehat{\jmath}}(u) = \|\varphi'_{\widehat{\jmath}}(u)\|, \tag{2.6}$$

а стационарность точки $\bar{u}$ в задаче (1.11) означает, что $\varphi'_{\widehat{\jmath}}(\bar{u}) = 0$. Полуустойчивость, а значит, и оценка расстояния (2.3), при этом сводятся к невырожденности матрицы $\varphi''_{\widehat{\jmath}}(\bar{u})$. Далее, поскольку для всякого $v \in \mathbb{R}^p$

$$\varphi''_{\widehat{\jmath}}(\bar{u})v = ((\Phi^{\widehat{\jmath}})'(\bar{u}))^{\top}(\Phi^{\widehat{\jmath}})'(\bar{u})v + ((\Phi^{\widehat{\jmath}})''(\bar{u})[v])^{\top}\Phi^{\widehat{\jmath}}(\bar{u}),$$

то

$$\langle \varphi''_{\widehat{\jmath}}(\bar{u})v,\, v \rangle = \|(\Phi^{\widehat{\jmath}})'(\bar{u})v\|^2 + \langle \Phi^{\widehat{\jmath}}(\bar{u}),\, (\Phi^{\widehat{\jmath}})''(\bar{u})[v,\, v] \rangle,$$

и, например, если квадратичная форма $v \mapsto \langle \Phi^{\widehat{\jmath}}(\bar{u}),\, (\Phi^{\widehat{\jmath}})''(\bar{u})[v,\, v]\rangle : \mathbb{R}^p \to \mathbb{R}$ положительно определена на $\ker(\Phi^{\widehat{\jmath}})'(\bar{u})$, то матрица $\varphi''_{\widehat{\jmath}}(\bar{u})$ также положительно определена, и, в частности, невырождена. Сказанное справедливо и для произвольного P, если $\bar{u} \in \operatorname{int} P$ (при этом равенство (2.6) имеет место для всех $u \in \mathbb{R}^p$, достаточно близких к $\bar{u}$).

В завершение этого обсуждения заметим, что характеризация (или хотя бы какие-то достаточные условия для) требуемой оценки расстояния (2.3) могут быть получены и в случае неполиэдрального P, однако, по-видимому, это требует задания P какими-то функциональными ограничениями и предположений о выполнении условий регулярности для этих ограничений, таких, как условие постоянного ранга. Такие предположения позволяют непосредственно вычислять производных по направлениям функции $r_{\widehat{\jmath}}$ [8, теорема 4.5.3], с последующим применением [7, предложение 1.64], согласно которому требуемая оценка расстояния равносильна тому, что все такие производные по ненулевым направлениям отличны от нуля.

Способ идентификации активных в точке $\bar{u}$ гладких кусочных отображений при выполнении оценки расстояния (2.3) описывается следующим результатом.

П р е д л о ж е н и е 2.1. *Пусть для некоторого* $\widehat{\jmath} \in \{1,\, \ldots,\, s\}$ *точка* $\bar{u}$ *является стационарной в задаче* (1.11), *и выполняется оценка расстояния* (2.3) *при* $u \in P$ *стремящемся к* $\bar{u}$. *Пусть* $\rho : \mathbb{R}_+ \to \mathbb{R}_+$ *— любая функция, удовлетворяющая* $\rho(0) = 0$, $\rho(t) \to 0$ *и* $t = o(\rho(t))$ *при* $t \to 0+$. *Для каждого* $u \in \mathbb{R}^p$ *положим*

$$\mathcal{A}_{\widehat{\jmath}}(u) = \{j \in \{1,\, \ldots,\, s\} \mid \|\Phi^j(u) - \Phi(u)\| \leqslant \rho(r_{\widehat{\jmath}}(u))\}. \tag{2.7}$$

Тогда для любого $u \in P$, *достаточно близкого к* $\bar{u}$, *выполняется*

$$\mathcal{A}_{\widehat{\jmath}}(u) = \mathcal{A}(\bar{u}).$$

Д о к а з а т е л ь с т в о. Пусть $j \in \mathcal{A}(\bar{u})$. Тогда, согласно (2.3), для $u \in P$, достаточно близкого к $\bar{u}$, выполняется

$$\|\Phi^j(u) - \Phi(u)\| = \|(\Phi^j(u) - \Phi(u)) - (\Phi^j(\bar{u}) - \Phi(\bar{u}))\| = O(\|u - \bar{u}\|) = O(r_{\widehat{\jmath}}(u)) \leqslant \rho(r_{\widehat{\jmath}}(u)),$$

и следовательно $j \in \mathcal{A}_{\widehat{\jmath}}(u)$.

Наоборот, если $j \notin \mathcal{A}(\bar{u})$, то $\|\Phi^j(\bar{u}) - \Phi(\bar{u})\| > 0$, и поскольку $r_{\widehat{\jmath}}(u) \to 0$ при $u \to \bar{u}$, отсюда следует, что $j \notin \mathcal{A}_{\widehat{\jmath}}(u)$ для u достаточно близкого к $\bar{u}$. □

3. Глобализованный кусочный метод Левенберга–Марквардта с процедурой выхода из окрестности нестационарной точки

Прототипом для предлагаемого ниже алгоритма 2 служит [5, алгоритм 4.1], реализующий глобализованный метод LP-Newton с процедурой выхода из стационарных точек, не являющихся решениями.

Для всякого $u \in \mathbb{R}^p$ введем обозначения $r_J(u) = r_{\widehat{\jmath}}(u)$ и $\mathcal{A}_J(u) = \mathcal{A}_{\widehat{\jmath}}(u)$ для $\widehat{\jmath} \in \mathcal{A}(u)$ такого, что $J(u) = (\Phi^{\widehat{\jmath}})'(u)$, т. е., согласно (2.1), (2.2), (2.7),

$$r_J(u) = \left\| u - \pi_P\left(u - (J(u))^\top \Phi(u)\right) \right\|. \tag{3.1}$$

$$\mathcal{A}_J(u) = \{j \in \{1,\, \ldots,\, s\} \mid \|\Phi^j(u) - \Phi(u)\| \leqslant \rho(r_J(u))\}. \tag{3.2}$$

Заметим, что при этом всегда выполняется $\mathcal{A}(u) \subset \mathcal{A}_J(u)$, и в частности, $\widehat{\jmath} \in \mathcal{A}_J(u)$.

Алгоритм 2. Фиксируем отображение $J: \mathbb{R}^p \to \mathbb{R}^{q \times p}$, удовлетворяющее (1.3). Фиксируем параметры $\theta > 0$, $\delta_0 > 0$, $\delta_1 > 0$, $\nu \in (0, 1)$, $\varepsilon \in (0, 1)$, и $\varkappa \in (0, 1)$. Выбираем функцию $\rho: \mathbb{R}_+ \to \mathbb{R}_+$, удовлетворяющую $\rho(t) \to 0$, $t = o(\rho(t))$ при $t \to 0+$. Выбираем $u^0 \in P$ и полагаем $k = 0$.

1. Если $\Phi(u^k) = 0$, стоп. Иначе вычисляем $\sigma(u^k)$ согласно (1.7).

2. Пусть индекс $\widehat{\jmath} \in \mathcal{A}(u^k)$ таков, что $J(u^k) = (\Phi^{\widehat{\jmath}})'(u^k)$. Если
$$\frac{(r_J(u^k))^\nu}{\|\Phi(u^k)\|} \leqslant \delta_0 \tag{3.3}$$
и $\mathcal{A}_J(u^k) \neq \{\widehat{\jmath}\}$, переходим к шагу 3. В противном случае, вычисляем v^k как решение подзадачи (1.4), и если $v^k \neq 0$, переходим к шагу 5; иначе стоп.

3. Выбираем $j \in \mathcal{A}_J(u^k) \setminus \{\widehat{\jmath}\}$ такой, что $r_j(u^k) \geqslant \delta_1$. Если такого j не существует, выбираем $j \in \mathcal{A}_J(u^k) \setminus \{\widehat{\jmath}\}$ с максимальным значением $r_j(u^k)$.

   Если $r_j(u^k) > r_J(u^k)$, вычисляем $v^{k,j}$ как решение подзадачи
$$\frac{1}{2}\|\Phi(u^k) + (\Phi^j)'(u^k)v\|^2 + \frac{1}{2}\sigma(u^k)\|v\|^2 \to \min, \quad u^k + v \in P, \tag{3.4}$$
и переходим к шагу 4. В противном случае, если $r_J(u^k) > 0$, вычисляем $v^{k,\widehat{\jmath}}$ как решение подзадачи (1.4), полагаем $v^k = v^{k,\widehat{\jmath}}$, и переходим к шагу 5; иначе стоп.

4. Полагаем $\alpha = 1$. Если выполняется неравенство
$$\varphi_j(u^k + \alpha v^{k,j}) \leqslant \varphi_j(u^k) - \varepsilon\alpha\sigma(u^k)\|v^{k,j}\|^2, \tag{3.5}$$
полагаем $\alpha_{k,j} = \alpha$. Иначе заменяем α на $\varkappa\alpha$ до тех пор, пока неравенство (3.5) не будет выполнено.

   Если
$$\varphi(u^k + \alpha_{k,j}v^{k,j}) < \varphi(u^k), \tag{3.6}$$
полагаем $v^k = v^{k,j}$, $\alpha_k = \alpha_{k,j}$, и переходим к шагу 6. В противном случае полагаем $v^k = v^{k,\widehat{\jmath}}$.

5. Полагаем $\alpha = 1$. Если выполняется неравенство
$$\varphi(u^k + \alpha v^k) \leqslant \varphi(u^k) - \varepsilon\alpha\sigma(u^k)\|v^k\|^2, \tag{3.7}$$
полагаем $\alpha_k = \alpha$. Иначе заменяем α на $\varkappa\alpha$ до тех пор, пока неравенство (3.7) не будет выполнено.

6. Полагаем $u^{k+1} = u^k + \alpha_k v^k$, увеличиваем k на 1, и переходим к шагу 1.

В сформулированном алгоритме подразумевается, что значения π_P, а значит, и r_j, относительно легко вычисляются. Если трудоемкость вычисления π_P сравнима с трудоемкостью решения подзадачи метода LM, то на шаге 3 алгоритма вместо $r_j(u^k)$ можно использовать $\|v^{k,j}\|$ для предварительно вычисленного $v^{k,j}$.

З а м е ч а н и е 3.1. Остановка алгоритма 2 для некоторого k происходит только в следующих случаях:

- если u^k является решением (1.1) (остановка на шаге 1);
- если $(r_J(u^k))^\nu/\|\Phi(u^k)\| > \delta_0$ (т. е. нарушается условие (3.3)) или $\mathcal{A}_J(u^k) = \{\widehat{\jmath}\}$, и $v^{k,\widehat{\jmath}} = 0$ (остановка на шаге 2); но последнее возможно только при $r_J(u^k) = 0$, и значит, согласно (2.7), $\mathcal{A}(u^k) = \mathcal{A}_J(u^k) = \{\widehat{\jmath}\}$, и при этом выполняется условие стационарности
$$\langle \varphi'_{\widehat{\jmath}}(u^k),\, u - u^k\rangle \geqslant 0 \quad \forall\, u \in P;$$
- если $r_j(u^k) \leqslant r_J(u^k) = 0$ (остановка на шаге 3); при этом не может выполняться $r_j(u^k) \geqslant \delta_1$, а значит, согласно выбору j на шаге 3, $r_j(u^k) = 0$ для всех $j \in \mathcal{A}_J(u^k)$, и поскольку при этом $\mathcal{A}_J(u^k) = \mathcal{A}(u^k)$, то выполняется
$$\langle \varphi'_j(u^k),\, u - u^k\rangle \geqslant 0 \quad \forall\, u \in P,\ \forall\, j \in \mathcal{A}(u^k). \tag{3.8}$$

З а м е ч а н и е 3.2. При переборе $j \in \mathcal{A}_J(u^k) \setminus \{\widehat{\jmath}\}$ на шаге 3 алгоритма 2 имеет смысл сначала пробовать индексы j с наименьшим значением $\|\Phi^j(u^k) - \Phi(u^k)\|$.

З а м е ч а н и е 3.3. Параметры δ_0 и δ_1 в алгоритме 2 предполагаются положительными и фиксированными, но можно рассматривать и какие-то динамические правила управления этими параметрами. Для справедливости излагаемой ниже теории достаточно предполагать, что эти параметры не становятся меньше некоторых фиксированных положительных величин.

Для обоснования глобальной сходимости алгоритма 2 потребуется следующая

Лемма 3.1. *Для любых $\bar{u} \in P$ и $j \in \mathcal{A}(\bar{u})$ таких, что $r_j(\bar{u}) > 0$, существует $\widetilde{\delta} > 0$ такое, что для любого $u^k \in P$, достаточно близкого к $\bar{u}$, подзадача (3.4) имеет единственное решение v^k, и для него выполняется $\|v^k\| \geqslant \widetilde{\delta}$.*

Доказательство этой леммы по сути содержится в доказательстве в [3, теорема 3.3]. Следующее предложение выводится как дополнительный факт, также вытекающий из доказательства в [3, теорема 3.3].

П р е д л о ж е н и е 3.1. *Пусть $\{u^k\} \subset P$ — любая последовательность, для которой последовательность $\{\|\Phi(u^k)\|\}$ монотонно невозрастает, и для любой предельной точки $\bar{u}$ последовательности $\{u^k\}$ существует сходящаяся к $\bar{u}$ подпоследовательность $\{u^{k_i}\}$ такая, что точка u^{k_i+1} получена шагом алгоритма 1 для всех i.*

Тогда для любой такой предельной точки выполняется (1.10) для некоторого $j \in \mathcal{A}(\bar{u})$, и $r_J(u^{k_i}) \to 0$ при $i \to \infty$.

Свойства глобальной сходимости алгоритма 2 характеризуются следующей теоремой.

Теорема 3.1. *Пусть выполнено условие (1.9).*

Тогда алгоритм 2 либо останавливается в точке $u^k \in P$, удовлетворяющей (3.8), либо генерирует бесконечную последовательность $\{u^k\} \subset P$. В последнем случае, если $\bar{u}$ — предельная точка этой последовательности (автоматически лежащая в P в силу его замкнутости), то либо оценка расстояния (2.3) при $u \in P$ стремящемся к $\bar{u}$ нарушается при некотором $\widehat{\jmath} \in \mathcal{A}(\bar{u})$, либо
$$\langle \varphi'_j(\bar{u}),\, u - \bar{u}\rangle \geqslant 0 \quad \forall\, u \in P,\ \forall\, j \in \mathcal{A}(\bar{u}). \tag{3.9}$$

Д о к а з а т е л ь с т в о. Если алгоритм 2 не останавливается в точке u^k, то очередное приближение u^{k+1} всегда успешно определяется алгоритмом. Это устанавливается теми же рассуждениями, что и в [3, теорема 3.3]. А именно, подзадачи (1.4) на шаге 2 и (3.4) на шаге 3 алгоритма всегда имеют имеют единственные решения, а процедуры одномерного поиска на шагах 4 и 5 алгоритма позволяют найти подходящие значения $\alpha_{k,j}$ и α_k, соответственно, после конечного числа дроблений, т. е. замен α на $\kappa\alpha$. Кроме того, согласно конструкции алгоритма, все генерируемые им приближения лежат в P.

Далее, согласно замечанию 3.1, с учетом (2.1), (2.2) получаем, что если алгоритм останавливается в точке u^k, то выполняется (3.8). Если же алгоритм не останавливается, то, согласно сказанному выше, он генерирует бесконечную последовательность $\{u^k\} \subset P$. Пусть $\bar{u} \in P$ — предельная точка этой последовательности, и предположим, что неравенство в (3.9) нарушается для некоторого $j \in \mathcal{A}(\bar{u})$. Последнее эквивалентно тому, что $r_j(\bar{u}) \neq 0$ и, в частности, в силу (1.2), (2.1) и (2.2), $\Phi(\bar{u}) = \Phi^j(\bar{u}) \neq 0$.

Пусть $\{u^{k_i}\}$ — сходящаяся к $\bar{u}$ подпоследовательность последовательности $\{u^k\}$, и предположим сначала, что для всех i на шаге 2 алгоритма реализуется

$$\frac{(r_J(u^{k_i}))^\nu}{\|\Phi(u^{k_i})\|} > \delta_0. \tag{3.10}$$

Тогда приближения u^{k_i+1} генерируются алгоритмом 1 (без каких-либо модификаций), и поэтому из предложения 3.1 вытекает, что $r_J(u^{k_i}) \to 0$ при $i \to \infty$, что противоречит (3.10). Таким образом, для любой сходящейся к $\bar{u}$ подпоследовательности $\{u^{k_i}\}$ последовательности $\{u^k\}$ модификация алгоритма 1 инициируется на итерации k_i для всех достаточно больших i, а именно, на шаге 2 алгоритма 2 выполняется $(r_J(u^{k_i}))^\nu/\|\Phi(u^{k_i})\| \leqslant \delta_0$, и если $\mathcal{A}_{\widehat{\jmath}}(u^k) \neq \{\widehat{\jmath}\}$, то осуществляется переход к шагу 3.

Предположим, что оценка расстояния (2.3) при $u \in P$ стремящемся к $\bar{u}$ выполняется для всех $j \in \mathcal{A}(\bar{u})$. Поскольку по непрерывности $\mathcal{A}(u^{k_i}) \subset \mathcal{A}(\bar{u})$ для всех достаточного больших i, то $\widehat{\jmath}$, отвечающий u^{k_i} при таких i, содержится в $\mathcal{A}(\bar{u})$, а значит, применимо предложение 2.1 с любым таким $\widehat{\jmath}$, позволяющее заключить, что $\mathcal{A}_J(u^{k_i}) = \mathcal{A}(\bar{u})$ для всех достаточного больших i.

Далее, из существования $j \in \mathcal{A}(\bar{u})$ такого, что $r_j(\bar{u}) \neq 0$, и из правила выбора j на шаге 3 алгоритма 2 вытекает существование $\widehat{\delta} > 0$ такого, что для всех таких j при достаточно большом i выполняется $r_j(u^{k_i}) \geqslant \widehat{\delta}$. В силу конечности множества $\mathcal{A}(\bar{u})$, еще раз переходя при необходимости к подпоследовательности, можем считать, что для $j = \widetilde{\jmath}$ при некотором фиксированном $\widetilde{\jmath} \in \mathcal{A}(\bar{u})$, и для всех i алгоритм использует либо $v^{k_i} = v^{k_i,\widetilde{\jmath}}$, вычисленное на шаге 3 как решение подзадачи (3.4), и удовлетворяющее (3.5) $\alpha^{k_i} = \alpha^{k_i,\widetilde{\jmath}}$, вычисленное на шаге 4, либо $\widetilde{\jmath} = \widehat{\jmath} \in \mathcal{A}(u^{k_i})$, и алгоритм использует v^{k_i}, вычисленное на шаге 3 как решение подзадачи (3.4), и удовлетворяющее (3.7) α_{k_i}, вычисленное на шаге 5, причем в любом случае $r_{\widetilde{\jmath}}(u^{k_i}) \geqslant \widehat{\delta}$.

Согласно лемме 3.1, из последнего неравенства следует существование $\widetilde{\delta} > 0$ такого, что для всех достаточно больших i выполняется $\|v^{k_i}\| \geqslant \widetilde{\delta}$, и при этом из (1.9), и (3.5) или (3.7), вытекает выполнение

$$\varphi(u^{k_i+1}) \leqslant \varphi_{\widetilde{\jmath}}(u^{k_i}) - \varepsilon\alpha_{k_i}\sigma(u^{k_i})\widetilde{\delta}^2. \tag{3.11}$$

(Напомним, что всегда $\varphi(u^{k_i}) = \varphi_{\widetilde{\jmath}}(u^{k_i})$.) Конструкция алгоритма такова, что последовательность $\{\varphi(u^k)\}$ монотонно невозрастает (в частности, для обеспечения этого предназначен тест (3.6)), причем эта последовательность ограничена снизу (нулем), а значит, сходится. Отсюда следует, что для (любой) предельной точки точки $\bar{u}$ последовательности $\{u^k\}$ последовательность $\{\varphi(u^k)\}$ (а значит, и ее подпоследовательнось $\{\varphi(u^{k_i+1})\}$) сходится к $\varphi(\bar{u})$, в силу непрерывности φ. С другой стороны, из сходимости $\{u^{k_i}\}$ к $\bar{u}$ и из непрерывности $\varphi_{\widetilde{\jmath}}$ следует, что $\{\varphi_{\widetilde{\jmath}}(u^{k_i})\} \to \varphi_{\widetilde{\jmath}}(\bar{u}) = \varphi(\bar{u})$, где последнее равенство имеет место потому, что $\widetilde{\jmath} \in \mathcal{A}(\bar{u})$). С учетом того, что $\sigma(\bar{u}) > 0$ (так как $\Phi(\bar{u}) \neq 0$), а значит, значения $\sigma(u^{k_i})$ отделены от нуля положительной константой, из (3.11) при этом следует, что $\alpha_{k_i} \to 0$ при $i \to \infty$. Но тогда, повторяя соответствующее рассуждение из [3, теорема 3.3] для $j = \widetilde{\jmath}$, снова приходим к противоречию. $\square$

Наконец, приведем результат о сверхлинейной скорости сходимости, демонстрирующий, что соответствующие свойства алгоритма 1, установленные в [3, теорема 3.4], сохраняются и для алгоритма 2.

Теорема 3.2. *Пусть производные кусочных отображениями $\Phi^1, \ldots, \Phi^s$ удовлетворяют условию Липшица вблизи решения $\bar{u}$ задачи* (1.1)*, а множество P является параллелепипедом, т. е. имеет вид* (2.5)*. Пусть выполнено P-свойство* (1.5) *с некоторой окрестностью U точки $\bar{u}$ (что является автоматическим при выполнении* (1.9)*), а также кусочная оценка расстояния* (1.6) *при $u \in P$ стремящемся к $\bar{u}$.*

Тогда если алгоритм 2*, в котором используется $\theta \in (0, 2]$, генерирует приближение достаточно близкое к $\bar{u}$, то алгоритм либо останавливается в решении u^k задачи* (1.1)*, либо генерирует бесконечную последовательность $\{u^k\}$, сходящуюся к некоторому решению $\bar{u}$ задачи* (1.1)*, причем скорость сходимости сверхлинейная с Q-порядком* $\min\{\theta + 1, 2\}$.

Доказательство. Напомним, что для индекса $\widehat{\jmath}$, отвечающего u^k, достаточно близкому к $\bar{u}$, выполняется $\widehat{\jmath} \in \mathcal{A}(u^k) \subset \mathcal{A}(\bar{u})$. Тогда, с учетом (2.2), из (1.6) и из [9, лемма 1] следует свойство верхней липшицевости решений вариационного неравенства (2.4). Точнее, для $w \in \mathbb{R}^p$ и для любого решения $u(w)$ возмущенного вариационного неравенства

$$u \in P, \quad \langle \varphi'_{\widehat{\jmath}}(u) - w, \widetilde{u} - u \rangle \geqslant 0 \quad \forall\, \widetilde{u} \in P,$$

достаточно близкого к $\bar{u}$, имеет место оценка

$$\operatorname{dist}(u(w), (\Phi^{\widehat{\jmath}})^{-1}(0) \cap P) = O(\|w\|)$$

при $w \to 0$. В частности, полагая $w = 0$, отсюда получаем, что вблизи $\bar{u}$ решения кусочной задачи

$$\Phi^{\widehat{\jmath}}(u) = 0, \quad u \in P,$$

совпадают с решениями вариационного неравенства (2.4).

Далее, для множества P из (2.5) вариационное неравенство (2.4) является так называемой смешанной комплементарной задачей, и для нее из установленного свойства верхней липшицевости и из [10, теорема 2] следует выполнение оценки расстояния

$$\operatorname{dist}(u, (\Phi^{\widehat{\jmath}})^{-1}(0) \cap P) = O(\|r_J(u)\|)$$

при $u \to \bar{u}$. (Отметим, что для случая $P = \mathbb{R}^n$ такая оценка вытекает непосредственно из [10, следствие 2].) В комбинации с (1.6) и условием $\nu \in (0, 1)$ эта оценка позволяет утверждать, что условие (3.3) нарушается для любого $u^k \in \mathbb{R}^p$, достаточно близкого к $\bar{u}$

и такого, что $\Phi(u^k) \neq 0$. Значит, в этом случае следующее приближение u^{k+1} вычисляется так же, как в алгоритме 1 (без каких-либо модификаций), и требуемый результат получается так же, как [3, теорема 3.4], с применением [3, теорема 2.1]. □

4. Иллюстративный пример

Следующий пример взят из [3, пример 3.1], [5, примеры 1.2, 2.1, 4.1].

П р и м е р 4.1. Пусть $p = 1$, $q = 2$,

$$\Phi(u) = \begin{pmatrix} 1-u \\ \min\{1+u,\, 1-u\} \end{pmatrix}, \quad P = [-1,\, 1].$$

Задача (1.1) с такими данными имеет единственное решение $\widehat{u} = 1$. Отображение Φ является кусочно-гладким, с естественным набором из $s = 2$ гладких кусочных отображений

$$\Phi^1(u) = \begin{pmatrix} 1-u \\ 1+u \end{pmatrix}, \quad \Phi^2(u) = \begin{pmatrix} 1-u \\ 1-u \end{pmatrix},$$

из которых активным в решении $\widehat{u}$ является только Φ^2, т. е. $\mathcal{A}(\widehat{u}) = \{2\}$. Отметим, что такая задача (1.1) является эквивалентной переформулировкой комплементарной системы

$$1-u = 0, \quad 1+u \geqslant 0, \quad 1-u \geqslant 0, \quad (1+u)(1-u) = 0,$$

и как и для всякой такой переформулировки, для нее выполняется условие (1.9); см. [3, пример 3.1].

Далее, непосредственно проверяется, что в точке $\bar{u} = 0 \in \operatorname{int} P$ выполняется $\mathcal{A}(\bar{u}) = \{1,\, 2\}$, причем $\varphi_1'(\bar{u}) = 0$, в то время как $\varphi_2'(\bar{u}) = -2$, т. е. точка $\bar{u}$ удовлетворяет (1.10) для $j = 1$, но не для $j = 2$. При этом, как показано в [3, пример 3.1], будучи запущен из точки $u^0 \in [-1,\, 0)$, алгоритм 1 генерирует монотонно возрастающую последовательность, сходящуюся к $\bar{u}$ линейно, с асимптотическим общим частным $2^{\theta/2}/(2+2^{\theta/2})$. Если же $u^0 \in (0,\, 1)$, то аналогичным анализом можно показать, что имеет место монотонная сходимость к решению $\widehat{u} = 1$, причем скорость сходимости сверхлинейная с Q-порядком $\theta + 1$.

Что же касается точки $u^0 = \bar{u} = 0$, то в этом случае поведение алгоритма определяется выбором $J(0):$ если выбирается $J(0) = (\Phi^1)'(0)$, то алгоритм 1 останавливается в этой точке, поскольку генерирует $v^k = 0$; если же выбирается $J(0) = (\Phi^2)'(0)$, то алгоритм «выходит» из точки u^0, генерируя $u^1 \in (0,\, 1)$, и имеет место сверхлинейная сходимость к решению $\widehat{u} = 1$. Идея алгоритма 2 как раз и состоит в переключении в подобных ситуациях на альтернативное гладкое кусочное отображение, что может давать возможность покинуть стационарную точку для текущей ветви, не являющуюся решением, или ее окрестность. Заметим, что $\varphi_1''(\bar{u}) = 2 \neq 0$, и выполнение оценки расстояния (2.3) при $\widehat{\jmath} = 1$ здесь гарантировано.

Пусть для определенности $J(0) = (\Phi^1)'(0)$, т. е.

$$J(u) = \begin{cases} (\Phi^1)'(u) = (-1,\, 1), & \text{если} \quad -1 \leqslant u \leqslant 0, \\ (\Phi^2)'(u) = (-1,\, -1), & \text{если} \quad 0 < u \leqslant 1. \end{cases}$$

Возьмем, например, $\rho(t) = \sqrt{t}$. Тогда, из (3.1) и (3.2) прямыми вычислениями выводится, что

$$\mathcal{A}_J(u) = \begin{cases} \{1\} = \mathcal{A}(u), & \text{если} \quad -1 \leqslant u < -1/2, \\ \{1,\, 2\}, & \text{если} \ -1/2 \leqslant u \leqslant \widetilde{u}, \\ \{2\} = \mathcal{A}(u), & \text{если} \quad \widetilde{u} < u \leqslant 1, \end{cases}$$

где $\widetilde{u} = (-1 + \sqrt{33})/8 \approx 0.5931$ — положительный корень уравнения $4u^2 + u - 2 = 0$.

Таким образом, если $-1 \leqslant u^k < -1/2$ или $\widetilde{u} < u^k < 1$, то альтернативные гладкие кусочные отображения не рассматриваются, и следующее приближение u^{k+1}, генерируемое алгоритмом 2, совпадает с приближением, получаемым алгоритмом 1. Во втором случае это имеет место и для всех последующих приближений, что приводит к сверхлинейной сходимости последовательности $\{u^k\}$ к решению $\widehat{u} = 1$. В первом же случае после конечного числа итераций очередное приближение будет удовлетворять $-1/2 \leqslant u^k \leqslant 0$, и при этом $\mathcal{A}_J(u^k) = \{1,\, 2\} \neq \{\widehat{\jmath}\} = \{1\}$, и

$$\frac{(r_J(u^k))^\nu}{\|\Phi(u^k)\|} = \frac{(-2u^k)^\nu}{\sqrt{2(1 + (u^k)^2)}},$$

где правая часть стремится к 0 при $u^k \to 0$. Поэтому, для любых фиксированных $\delta_0 > 0$ и $\nu > 0$, на шаге 2 алгоритма 2 тест (3.3) будет выполнен либо для текущего u^k, либо для какого-то из последующих приближений u^k, когда оно станет достаточно близким в 0. Тогда на шаге 3 алгоритма 2 вычисляется величина $r_2(u^k) = 2 - u^k > r_J(u^k) = -2u^k$ при u^k достаточно близком к 0. Поэтому следующее приближение u^{k+1} определяется шагами 3, 4 и 6 алгоритма 2 с $j = 2$, и прямые вычисления показывают, что при этом всегда $u^{k+1} > u^k$, и $u^{k+1} > 0$ или $u^{k+2} > 0$ (в зависимости от того, насколько близко к 0 расположено u^k).

Остается рассмотреть случай, когда $0 < u^k \leqslant \widetilde{u}$. При этом $\mathcal{A}_J(u^k) = \{1,\, 2\} \neq \{\widehat{\jmath}\} = \{2\}$ и

$$\frac{(r_J(u^k))^\nu}{\|\Phi(u^k)\|} = \frac{(2 - u^k)^\nu}{1 - u^k}.$$

Функция в правой части относительно u^k монотонно возрастает на $[0,\, 1)$, а значит ее значения всюду в этой области не меньше, чем значение в 0, равное 2^ν. Соответственно, если, например, $\delta_0 \leq 2^\nu$, то тест (3.3) не выполняется, и все последующие итерации осуществляются шагами алгоритма 1, что приводит к сверхлинейной сходимости к решению $\widehat{u} = 1$.

Если же $\delta_0 > 2^\nu$, то для u^k близких к 0 тест (3.3) выполняется, и тогда на шаге 3 алгоритма 2 вычисляется величина $r_1(u^k) = -2u^k < r_J(u^k) = 2 - u^k$. Поэтому, независимо от выполнения условия $r_1(u^k) \geqslant \delta_1$, следующее приближение u^{k+1} определяется шагами 5 и 6 алгоритма 1. При этом $u^{k+1} > u^k$, откуда следует, что и все последующие итерации осуществляются шагами алгоритма 1, что снова приводит к сверхлинейной сходимости к решению.

References

[1] A. Fischer, A. F. Izmailov, M. V. Solodov, “The Levenberg–Marquardt method: an overview of modern convergence theories and more”, *Computational Optimization and Applications*, **89**:1 (2024), 33–67.

[2] A. F. Izmailov, E. I. Uskov, Z. Yan, "The piecewise Levenberg–Marquardt method", *Advances in System Sciences and Applications*, **24**:1 (2024), 29–39.

[3] A. F. Izmailov, E. I. Uskov, Z. Yan, "Globalization of convergence of the constrained piecewise Levenberg–Marquardt method", *Optimization Methods and Software*, **40**:2 (2025), 243–265.

[4] A. Fischer, M. Herrich, A. F. Izmailov, M. V. Solodov, "Convergence conditions for Newton-type methods applied to complementarity systems with nonisolated solutions", *Computational Optimization and Applications*, **63**:2 (2016), 425–459.

[5] A. Fischer, M. Herrich, A. F. Izmailov, W. Scheck, M. V. Solodov, "A globally convergent LP-Newton method for piecewise smooth constrained equations: escaping nonstationary accumulation points", *Computational Optimization and Applications*, **69**:2 (2018), 325–349.

[6] A. Fischer, M. Herrich, A. F. Izmailov, M. V. Solodov, "A globally convergent LP-Newton method", *SIAM J. on Optimization*, **26** (2016), 2012–2033.

[7] A. F. Izmailov, M. V. Solodov, *Newton-Type Methods for Optimization and Variational Problems*, Springer Series in Operations Research and Financial Engineering, Springer, Cham, 2014.

[8] F. Facchinei, J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems.*, Springer–Verlag, New York, 2003.

[9] R. Behling, A. Fischer, "A unified local convergence analysis of inexact constrained Levenberg–Marquardt methods", *Optimization Letters*, **6** (2012), 927–940.

[10] A. Fischer, "Local behavior of an iterative framework for generalized equations with nonisolated solutions", *Mathematical Programming*, **94** (2002), 91–124.

Информация об авторах

Измаилов Алексей Феридович, доктор физико-математических наук, профессор кафедры исследования операций, Московский государственный университет им. М. В. Ломоносова, г. Москва, Российская Федерация. E-mail: izmaf@cs.msu.ru
ORCID: https://orcid.org/0000-0001-9851-0524

Янь Чжибай, аспирант, Московский государственный университет им. М. В. Ломоносова, г. Москва, Российская Федерация. E-mail: yanzhibai@cs.msu.ru
ORCID: https://orcid.org/0009-0003-6425-0332

Конфликт интересов отсутствует.

Для контактов:
Измаилов Алексей Феридович
E-mail: izmaf@cs.msu.ru

Поступила в редакцию 26.10.2025 г.
Поступила после рецензирования 19.11.2025 г.
Принята к публикации 21.11.2025 г.

Information about the authors

Alexey F. Izmailov, Doctor of Physical and Mathematical Sciences, Professor of the Operations Research Department, Lomonosov Moscow State University, Moscow, Russian Federation. E-mail: izmaf@cs.msu.ru
ORCID: https://orcid.org/0000-0001-9851-0524

Zhibai Yan, Post-Graduate Student, Lomonosov Moscow State University, Moscow, Russian Federation. E-mail: yanzhibai@cs.msu.ru
ORCID: https://orcid.org/0009-0003-6425-0332

There is no conflict of interests.

Corresponding author:
Alexey F. Izmailov
E-mail: izmaf@cs.msu.ru

Received 26.10.2025
Reviewed 19.11.2025
Accepted for press 21.11.2025