

УДК 519.626

О ВЛИЯНИИ КРИТИЧЕСКИХ МНОЖИТЕЛЕЙ ЛАГРАНЖА НА СКОРОСТЬ СХОДИМОСТИ МЕТОДА МНОЖИТЕЛЕЙ¹⁾

© 2012 г. А. Ф. Измаилов, Е. И. Усков

(119992 Москва, Ленинские горы, МГУ, ф-т ВМиК)

e-mail: izmaff@ccas.ru, ydoom@narod.ru

Поступила в редакцию 30.05.2012 г.

Статья посвящена исследованию влияния критических множителей Лагранжа на скорость сходимости метода множителей, а также на эффективность различных способов ускорения финальной фазы этого метода. Библ. 26. Фиг. 4. Табл. 3.

Ключевые слова: задача математического программирования, модифицированная функция Лагранжа, метод множителей, сверхлинейная сходимость, критический множитель Лагранжа, стабилизированный метод Ньютона–Лагранжа.

1. ВВЕДЕНИЕ. КРИТИЧЕСКИЕ МНОЖИТЕЛИ ЛАГРАНЖА

Рассматриваем задачу математического программирования

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad g(x) \leq 0, \quad (1.1)$$

где $f: \mathbb{R}^n \rightarrow \mathbb{R}$ — дважды дифференцируемая функция, а $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ — дважды дифференцируемые отображения.

Напомним, что точка $\bar{x} \in \mathbb{R}^n$ называется стационарной точкой задачи (1.1), если существуют такие $\bar{\lambda} \in \mathbb{R}^l$ и $\bar{\mu} \in \mathbb{R}^m$, что тройка $(\bar{x}, \bar{\lambda}, \bar{\mu})$ удовлетворяет системе Каруша–Куна–Таккера (ККТ)

$$\frac{\partial L}{\partial x}(x, \lambda, \mu) = 0, \quad h(x) = 0, \quad \mu \geq 0, \quad g(x) \leq 0, \quad \langle \mu, g(x) \rangle = 0 \quad (1.2)$$

этой задачи. Здесь $L: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}$ — функция Лагранжа задачи (1.1):

$$L(x, \lambda, \mu) = f(x) + \langle \lambda, h(x) \rangle + \langle \mu, g(x) \rangle.$$

При этом пара $(\bar{\lambda}, \bar{\mu})$ называется множителем Лагранжа, отвечающим стационарной точке $\bar{x}$. Множество таких пар будем обозначать $\mathcal{M}(\bar{x})$.

Определим $A(\bar{x}) = \{i = 1, \dots, m | g_i(\bar{x}) = 0\}$ — множество номеров ограничений-неравенств задачи (1.1), активных в допустимой точке $\bar{x}$ этой задачи. Введем также разбиение множества $A(\bar{x})$ на подмножества

$$A_+(\bar{x}, \bar{\mu}) = \{i \in A(\bar{x}) | \bar{\mu}_i > 0\}, \quad A_0(\bar{x}, \bar{\mu}) = \{i \in A(\bar{x}) | \bar{\mu}_i = 0\}.$$

Кроме того, положим $N(\bar{x}) = \{1, \dots, m\} \setminus A(\bar{x})$.

Используемое ниже достаточное условие второго порядка оптимальности в стационарной точке $\bar{x}$ задачи (1.1) для отвечающего ей множителя Лагранжа $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ имеет вид

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}, \bar{\mu}) \xi, \xi \right\rangle > 0 \quad \forall \xi \in C(\bar{x}) \setminus \{0\}, \quad (1.3)$$

где

$$\begin{aligned} C(\bar{x}) &= \{ \xi \in \mathbb{R}^n | h'(\bar{x})\xi = 0, g'_{A_0(\bar{x}, \bar{\mu})}(\bar{x})\xi \leq 0, \langle f'(\bar{x}), \xi \rangle \leq 0 \} = \\ &= \{ \xi \in \mathbb{R}^n | h'(\bar{x})\xi = 0, g'_{A_+(\bar{x}, \bar{\mu})}(\bar{x})\xi = 0, g'_{A_0(\bar{x}, \bar{\mu})}(\bar{x})\xi \leq 0 \} \end{aligned}$$

¹⁾Работа выполнена при финансовой поддержке РФФИ (код проекта 10-01-00251).

есть критический конус задачи (1.1) в точке $\bar{x}$. Здесь и далее через z_I обозначается подвектор вектора z с компонентами $z_i, i \in I$, где I – заданное конечное множество. Напомним, что в случае выполнения условия (1.3) точка $\bar{x}$ является строгим локальным решением задачи (1.1).

Ключевую роль в данной работе играет следующее понятие критического множителя Лагранжа, введенное в [1]: множитель $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ называется *критическим*, если существует набор $(\xi, \eta, \zeta) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ такой, что $\xi \neq 0$, и выполнены соотношения

$$\begin{aligned} \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}, \bar{\mu})\xi + (h'(\bar{x}))^T \eta + (g'(\bar{x}))^T \zeta &= 0, \quad h'(\bar{x})\xi = 0, \quad g'_{A_+(\bar{x}, \bar{\mu})}(\bar{x})\xi = 0, \\ \zeta_{A_0(\bar{x}, \bar{\mu})} &\geq 0, \quad g'_{A_0(\bar{x}, \bar{\mu})}(\bar{x})\xi \leq 0, \quad \zeta_i \langle g'_i(\bar{x}), \xi \rangle = 0, \quad i \in A_0(\bar{x}, \bar{\mu}), \quad \zeta_{N(\bar{x})} = 0, \end{aligned} \quad (1.4)$$

и *некритическим* иначе. Легко убедиться в том, что если в точке $\bar{x}$ для $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ выполнено достаточное условие второго порядка (1.3), то множитель $(\bar{\lambda}, \bar{\mu})$ является некритическим.

При отсутствии ограничений-неравенств из первых двух равенств в (1.4) вытекает, что критичность множителя $\bar{\lambda} \in \mathcal{M}(\bar{x})$ означает существование элемента $\xi \in \ker h'(\bar{x}) \setminus \{0\}$ такого, что

$$\frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi \in \text{im}(h'(\bar{x}))^T. \quad (1.5)$$

Для таких задач понятие критического множителя было введено в [2]. В случае полного вырождения (т.е. при $h'(\bar{x}) = 0$) критичность $\bar{\lambda}$ означает вырожденность матрицы Гессе $\frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})$. В общем случае, поскольку $\text{im}(h'(\bar{x}))^T = (\ker h'(\bar{x}))^\perp$, условие (1.5) эквивалентно условию

$$\pi_{\ker h'(\bar{x})} \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi = 0,$$

где через $\pi_{\ker h'(\bar{x})}$ обозначен оператор ортогонального проектирования на $\ker h'(\bar{x})$. Симметричную матрицу линейного оператора $\xi \rightarrow \pi_{\ker h'(\bar{x})} \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda})\xi : \ker h'(\bar{x}) \rightarrow \ker h'(\bar{x})$ называют

суженной матрицей Гессе функции Лагранжа по x . Критичность $\bar{\lambda}$ означает вырожденность этой матрицы.

Далее, из (1.2) немедленно следует, что если для множителя $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ выполняется условие $\bar{\mu}_{A(\bar{x}) \setminus A} = 0$ для некоторого множества индексов $A \subset A(\bar{x})$, то $\bar{x}$ является стационарной точкой задачи с ограничениями-равенствами

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad g_A(x) = 0, \quad (1.6)$$

в то время как $(\bar{\lambda}, \bar{\mu}_A)$ является отвечающим этой стационарной точке множителем Лагранжа. Понятие критического множителя Лагранжа дополняется следующим (введенным в [3]): множитель $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ называется *критическим относительно множества индексов $A \subset A(\bar{x})$* , если $\bar{\mu}_{A(\bar{x}) \setminus A} = 0$ и $(\bar{\lambda}, \bar{\mu}_A)$ – критический множитель Лагранжа, отвечающий стационарной точке $\bar{x}$ задачи (1.6). В противном случае $(\bar{\lambda}, \bar{\mu})$ называется *некритическим относительно A* .

Как легко видеть из приведенных определений, если множитель $(\bar{\lambda}, \bar{\mu})$ критический, то существует $I_1 \subset A_0(\bar{x}, \bar{\mu})$ такое, что этот множитель является критическим относительно множества индексов $A = A_+(\bar{x}, \bar{\mu}) \cup I_1$. В частности, если выполнено условие строгой дополнительнойности $\bar{\mu}_{A(\bar{x})} > 0$ (т.е. $A_0(\bar{x}, \bar{\mu}) = \emptyset$), то некритичность эквивалентна некритичности относительно $A(\bar{x})$.

Очевидно, что из выполнения достаточного условия второго порядка (1.3) вытекает выполнение соответствующего достаточного условия второго порядка оптимальности в стационарной точке $\bar{x}$ задачи (1.6) при $A = A(\bar{x})$ для отвечающего этой стационарной точке множителя Лагранжа

жа $(\bar{\lambda}, \bar{\mu}_{A(\bar{x})})$, и поэтому $(\bar{\lambda}, \bar{\mu})$ не может быть критическим множителем Лагранжа для задачи (1.1) относительно $A(\bar{x})$. Вместе с тем, критичность такого множителя относительно более узких множеств индексов $A \subset A(\bar{x})$ может иметь место (см. [3, пример 3.5]).

Одним из традиционных подходов к решению задачи (1.1) является метод модифицированных функций Лагранжа (или просто метод множителей, см. [4], [5]), продолжающий активно развиваться и по сей день (см. также [6], [7]). Данный метод является основой весьма успешных солверов, таких как LANCELOT (см. [8]) и ALGENCAN (см. [9]). Метод множителей обладает рядом привлекательных свойств (см. [10]), одно из которых заключается в сильной теории глобальной сходимости: стационарность предельных точек траекторий метода обосновывается при весьма слабых требованиях регулярности ограничений (см. [11], [12]), либо вовсе при отсутствии традиционной регулярности (см. [13]), и, в частности, при возможной неограниченности множества $\mathcal{M}(\bar{x})$, что делает методы множителей перспективной глобальной стратегией для задач оптимизации с нерегулярными ограничениями.

Другой привлекательной особенностью метода множителей является обоснованная локальная линейная и даже сверхлинейная сходимость без каких-либо условий регулярности. А именно, как показано в [14] (см. также теорему 1 ниже), если начальное приближение выбирается вблизи точки $(\bar{x}, \bar{\lambda}, \bar{\mu})$, где для $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ выполнено достаточное условие второго порядка оптимальности (1.3), то метод такого типа генерирует траекторию $\{(x^k, \lambda^k, \mu^k)\}$, которая сходится к точке $(\bar{x}, \lambda^*, \mu^*)$ при некотором $(\lambda^*, \mu^*) \in \mathcal{M}(\bar{x})$, близком к $(\bar{\lambda}, \bar{\mu})$. При этом скорость сходимости является линейной, а если параметр штрафа стремится к бесконечности, то и сверхлинейной, что согласуется с результатами о скорости сходимости метода множителей в более традиционных (и значительно более сильных) предположениях (см. [6, теорема 2.7], [7, теорема 4.7.4]). В частности, методам множителей присуще свойство двойственной стабилизации: двойственные траектории этих методов обычно сходятся даже в случае неограниченного множества множителей Лагранжа. Сказанное свидетельствует о том, что методы множителей могут в целом рассматриваться как перспективный подход к численному решению потенциально нерегулярных задач оптимизации, совмещающий глобальную сходимость с высокой скоростью сходимости.

Однако при ближайшем рассмотрении ситуация оказывается не столь однозначно обнадеживающей. Известно (см. [2], [3], [15], [16]), что для ньютоновских методов, применяемых к задачам оптимизации с нерегулярными ограничениями, имеет место эффект притяжения двойственных траекторий к критическим относительно некоторого множества индексов $A \subset A(\bar{x})$ множителям Лагранжа: такие множители и, в частности, те из них, для которых нарушено достаточное условие второго порядка, часто служат аттракторами для двойственных траекторий, и именно этот феномен обычно разрушает сверхлинейную скорость сходимости таких методов. Как показано в [17], эффект притяжения к критическим множителям Лагранжа наблюдается и для методов множителей. В настоящей работе изучается вопрос о том, насколько типичным является данный эффект для методов этого класса, и какое влияние он оказывает на скорость их сходимости.

Другая известная проблема, связанная с методами множителей, состоит в том, что увеличение параметра штрафа, необходимое для достижения сверхлинейной скорости сходимости, приводит к ухудшению обусловленности подзадач метода, что повышает вычислительные затраты на их решение. Высокая скорость сходимости может не компенсировать эти затраты, и метод со сверхлинейной скоростью сходимости может уступать по общей эффективности методу, в котором последовательность параметров штрафа остается ограниченной.

Возможное решение данной проблемы состоит в том, чтобы в финальной фазе метода (когда есть основания считать, что текущее приближение уже достаточно близко к решению) переключаться на некоторый локальный быстрый алгоритм. Данный подход к ускорению методов множителей, предложенный в [10], реализован в текущей версии солвера ALGENCAN: при достаточной малости невязки системы ККТ (1.2) в текущей точке осуществляется оценка множества активных в искомом решении ограничений-неравенств, формируется соответствующая задача с ограничениями-равенствами, и к ее системе уравнений Лагранжа применяется метод Ньютона. Проблема, однако, в том, что, как отмечено выше, в случае нерегулярных ограничений сверхлинейная скорость сходимости такого метода Ньютона–Лагранжа обычно все равно теряется из-за притяжения его двойственной траектории к критическим множителям Лагранжа.

Напрашивающейся альтернативой является использование в качестве ускорителя не метода Ньютона–Лагранжа, а его стабилизированной версии, изначально предложенной в [18] для за-

дач с ограничениями-неравенствами. В [19] приводится обзор дальнейших работ, посвященных этому методу, а в [1] для задачи с ограничениями-равенствами получен следующий результат о локальной сходимости стабилизированного метода Ньютона—Лагранжа: если начальное приближение близко к точке $(\bar{x}, \bar{\lambda})$, где $\bar{\lambda}$ — некритический множитель Лагранжа, отвечающий стационарной точке $\bar{x}$, то метод генерирует траекторию $\{(x^k, \lambda^k)\}$, которая сверхлинейно сходится к $(\bar{x}, \lambda^*)$ при некотором $\lambda^* \in \mathcal{M}(\bar{x})$, близком к $\bar{\lambda}$. В частности, тенденция притяжения к критическим множителям Лагранжа для этого метода не столь выражена, что отмечалось в [16]: существуют широкие области двойственных точек, при старте из которых такое притяжение для данного метода не имеет места. Вопрос, разумеется, в том, будут ли эти привлекательные свойства стабилизированного метода Ньютона—Лагранжа достаточными для его использования в качестве ускорителя в методах множителей: если двойственные траектории последнего притягиваются к критическим множителям, то точки этих траекторий могут никогда не попадать в область начальных приближений, подходящих для успешного применения такого ускорителя. Таким образом, и здесь роль критических множителей Лагранжа оказывается определяющей. Сравнительная эффективность двух указанных альтернативных ускорителей в случае нерегулярных ограничений также является предметом настоящей работы.

2. МЕТОД МНОЖИТЕЛЕЙ И ЕГО УСКОРИТЕЛИ

Определим семейство модифицированных функций Лагранжа задачи (1.1): $L_c: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}$,

$$L_c(x, \lambda, \mu) = f(x) + \frac{1}{2c} (\|\lambda + ch(x)\|_2^2 + \|\max\{0, \mu + cg(x)\}\|_2^2),$$

где $c > 0$ — параметр штрафа, а максимум берется покомпонентно. Метод множителей состоит в последовательном решении задач безусловной минимизации функции $L_c(\cdot, \lambda, \mu)$ при фиксированных значениях λ и μ , которые также итеративно пересчитываются. Отметим, что если среди ограничений задачи (1.1) присутствуют простые ограничения (например, нижние и верхние границы на компоненты x), то их часто не включают в модифицированную функцию Лагранжа, а учитывают непосредственно. В таком случае подзадача метода множителей состоит в минимизации $L_c(\cdot, \lambda, \mu)$ при простых прямых ограничениях (см. [11]). В частности, именно такой подход реализован в ALGENCAN (см. [9]). Здесь, исключительно для простоты изложения, будем считать, что прямых ограничений нет. С учетом этого упрощения, приводимый алгоритм метода множителей соответствует тому, который лежит в основе солвера ALGENCAN (см. [11, алгоритм 3.1], а также некоторые совсем недавние модификации и улучшения в [12], [20]).

Алгоритм 1.

Выбираем скалярные параметры $\bar{\lambda}_{\min}, \bar{\lambda}_{\max}, \bar{\lambda}_{\min} \leq \bar{\lambda}_{\max}, \bar{\mu}_{\max} \geq 0, c_0 > 0, \theta \in [0, 1)$ и $\rho > 1$. Полагаем $k = 0$.

Шаг 1. Выбираем вектор $(\bar{\lambda}^k, \bar{\mu}^k) \in \mathbb{R}^l \times \mathbb{R}^m$, удовлетворяющий $\bar{\lambda}_{\min} \leq \bar{\lambda}_j^k \leq \bar{\lambda}_{\max}, j = 1, \dots, l$, и $0 \leq \bar{\mu}_j^k \leq \bar{\mu}_{\max}, j = 1, \dots, m$. (Для $k \geq 1$ естественным является выбор $(\bar{\lambda}^k, \bar{\mu}^k)$ как проекции вектора (λ^k, μ^k) , получаемого в п. 2 алгоритма, на множество $\bigotimes_{j=1}^l [\bar{\lambda}_{\min}, \bar{\lambda}_{\max}] \times \bigotimes_{j=1}^m [0, \bar{\mu}_{\max}]$.)

Выбираем $\varepsilon_k \geq 0$ и вычисляем $x^{k+1} \in \mathbb{R}^n$ как приближенное решение задачи

$$L_{c_k}(x, \bar{\lambda}^k, \bar{\mu}^k) \rightarrow \min, \quad x \in \mathbb{R}^n, \quad (2.1)$$

а именно, как точку, удовлетворяющую условию

$$\left\| \frac{\partial L_{c_k}}{\partial x}(x^{k+1}, \bar{\lambda}^k, \bar{\mu}^k) \right\|_{\infty} \leq \varepsilon_k.$$

Шаг 2. Полагаем

$$\lambda^{k+1} = \bar{\lambda}^k + c_k h(x^{k+1}), \quad \mu^{k+1} = \max\{0, \bar{\mu}^k + c_k g(x^{k+1})\}.$$

Шаг 3. Если $k = 0$ или выполнено условие

$$\|(h(x^{k+1}), \min\{\mu^{k+1}, -g(x^{k+1})\})\|_{\infty} \leq \|(h(x^k), \min\{\mu^k, -g(x^k)\})\|_{\infty}, \quad (2.2)$$

выбираем любое значение $c_{k+1} \geq c_k$. В противном случае выбираем $c_{k+1} \geq \rho c_k$. Увеличиваем номер шага k на 1 и переходим к шагу 1.

В ALGENCAN используются следующие значения параметров: $\bar{\lambda}_{\min} = -10^{20}$, $\bar{\lambda}_{\max} = \bar{\mu}_{\max} = 10^{20}$, $\theta = 0.5$ и $\rho = 10$. Начальное значение параметра штрафа c_0 определяется с помощью специальной процедуры инициализации и зависит от значений f , h и g в имеющейся точке (см. [21]). Последующие значения параметра штрафа на шаге 3 алгоритма выбираются следующим образом: $c_{k+1} = c_k$, если условие (2.2) выполнено, и $c_{k+1} = \rho c_k$ в противном случае. Заметим, однако, что алгоритм 1 и теория его глобальной сходимости (см. [12]) допускает произвольное увеличение параметра штрафа при выполнении условия (2.2), а также более быстрое его увеличение при нарушении условия (2.2).

Отметим также, что помимо процедуры, описанной в алгоритме 1, ALGENCAN использует ряд эвристик, направленных на повышение его эффективности (см. [21]). В частности, используются “рестарты” для параметра штрафа: если текущее приближение оказывается “почти допустимым”, но при этом невязка остальных условий в системе ККТ (1.2) не является достаточно малой, то очередное значение параметра штрафа снова вычисляется с помощью процедуры инициализации. Кроме того, в ALGENCAN присутствует процедура масштабирования данных задачи: алгоритм 1 применяется не к исходной задаче (1.1), а к ее масштабированному варианту

$$\tilde{f}(x) \rightarrow \min, \quad \tilde{h}(x) = 0, \quad \tilde{g}(x) \leq 0, \quad (2.3)$$

где $\tilde{f}(x) = f(x)/\max\{1, \|f'(x^0)\|_\infty\}$, $\tilde{h}_i(x) = h_i(x)/\max\{1, \|h'_i(x_0)\|_\infty\}$, $i \in 1, \dots, l$, и $\tilde{g}_i(x) = g_i(x)/\max\{1, \|g'_i(x^0)\|_\infty\}$, $i \in 1, \dots, m$, а $x^0 \in \mathbb{R}^n$ — начальное приближение. Разумеется, переход к задаче (2.3) приводит к изменению как прямой, так и двойственной траектории метода. Вместе с тем, очевидно, что точка $\bar{x}$ является стационарной в исходной задаче (1.1) тогда и только тогда, когда она является стационарной в масштабированной задаче (2.3), хотя соответствующие множители Лагранжа могут меняться.

Локальная сходимость и скорость сходимости алгоритма 1 характеризуются следующей теоремой, доказанной в [14]. Введем функцию невязки $\sigma: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R} \rightarrow \mathbb{R}_+$ системы ККТ (1.2) следующим образом:

$$\sigma(x, \lambda, \mu) = \left\| \left(\frac{\partial L}{\partial x}(x, \lambda, \mu), h(x), \min\{\mu, -g(x)\} \right) \right\|. \quad (2.4)$$

Теорема 1. Пусть функция $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и отображения $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, и их вторые производные непрерывны в этой точке. Пусть $\bar{x}$ — стационарная точка задачи (1.1), удовлетворяющая достаточному условию второго порядка (1.3) при некотором $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$. Пусть $\varepsilon: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ — произвольная функция, удовлетворяющая условию $\varepsilon(t) = o(t)$.

Тогда для всякого $M > 0$ найдется $\bar{c} > 0$ такое, что для любого $c_0 \geq \bar{c}$ и любого начального приближения $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, достаточно близкого к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, существует удовлетворяющая для каждого $k = 0, 1, \dots$ условию

$$\|(x^{k+1} - x^k, \lambda^{k+1} - \lambda^k, \mu^{k+1} - \mu^k)\| \leq M\sigma(x^k, \lambda^k, \mu^k)$$

траектория $\{(x^k, \lambda^k, \mu^k)\}$ алгоритма 1, в котором $\varepsilon_k = \varepsilon(\sigma(x^k, \lambda^k, \mu^k))$; всякая такая траектория сходится к $(\bar{x}, \lambda^*, \mu^*)$ при некотором $(\lambda^*, \mu^*) \in \mathcal{M}(\bar{x})$, причем скорость сходимости линейная, и сверхлинейная, если $c_k \rightarrow +\infty$.

Теперь обсудим предложенную в [10] конструкцию ускорителя алгоритма 1, которая с незначительными изменениями реализована в ALGENCAN. При заданной точности остановки $\varepsilon > 0$, ускоритель запускается после того, как алгоритмом 1 достигается заданная точность $\hat{\varepsilon} > \varepsilon$, а именно, когда для текущего приближения (x^k, λ^k, μ^k) оказывается выполненным условие

$$\sigma(x^k, \lambda^k, \mu^k) \leq \hat{\varepsilon} \quad (2.5)$$

(в ALGENCAN в (2.4) используется ∞ -норма), но при этом еще не достигнута точность остановки, т.е. (2.5) не выполняется с $\hat{\varepsilon}$, замененным на ε ; в противном случае алгоритм останавливается, поскольку приближение (x^k, λ^k, μ^k) нужного качества найдено самим методом множителей.

Если запуск ускорителя оканчивается неудачей, то все его итерации отбрасываются, $\hat{\varepsilon}$ уменьшается (оставаясь большим, чем ε), и продолжается выполнение алгоритма 1 с того места, где оно было прервано, и до тех пор, пока не будет достигнута новая точность $\hat{\varepsilon}$. После этого вновь запускается ускоритель и т.д. При этом под неудачей ускорителя понимается ситуация, когда после заданного количества K его шагов не удастся получить приближение $(\tilde{x}, \tilde{\lambda}, \tilde{\mu}) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, удовлетворяющее условию

$$\sigma(\tilde{x}, \tilde{\lambda}, \tilde{\mu}) \leq \varepsilon. \quad (2.6)$$

Запуск ускорителя из точки $(\hat{x}, \hat{\lambda}, \hat{\mu}) = (x^k, \lambda^k, \mu^k)$ осуществляется следующим образом. Сначала формируется рабочее множество индексов $\hat{A} \subset \{1, \dots, m\}$, оценивающее множество номеров ограничений-неравенств активных в искомом решении:

$$\hat{A} = \{i = 1, \dots, m | g_i(\hat{x}) \geq -\hat{\varepsilon}\}. \quad (2.7)$$

Ограничения-неравенства с номерами $i \in \{1, \dots, m\} \setminus \hat{A}$ отбрасываются, а остальные ограничения-неравенства заменяются на равенства

$$g_i(x) + \frac{1}{2}s_i^2 = 0, \quad i \in \hat{A},$$

где $s_i, i \in \hat{A}$, — вспомогательные переменные. Согласно [10], введение вспомогательных переменных значительно повышает робастность ускорителя, снимая негативный эффект от возможного включения в $\hat{A}$ индексов ограничений, не являющихся активными в решении.

Далее, не ограничивая общности, предположим, что $\hat{A} = \{1, \dots, \hat{m}\}$ при некотором $\hat{m} \in \{0, 1, \dots, m\}$. Вводится отображение $\hat{h} : \mathbb{R}^n \times \mathbb{R}^{\hat{m}} \rightarrow \mathbb{R}^l \times \mathbb{R}^{\hat{m}}$,

$$\hat{h}(x, s) = \left(h(x), g_1(x) + \frac{1}{2}s_1^2, \dots, g_{\hat{m}}(x) + \frac{1}{2}s_{\hat{m}}^2 \right), \quad (2.8)$$

и к задаче

$$f(x) \rightarrow \min, \quad \hat{h}(x, s) = 0, \quad (2.9)$$

применяется метод Ньютона–Лагранжа (МНЛ), т.е. метод Ньютона для системы Лагранжа

$$\frac{\partial \hat{L}}{\partial(x, s)}((x, s), (\lambda, v)) = 0, \quad \hat{h}(x, s) = 0, \quad (2.10)$$

этой задачи, где $\hat{L} : (\mathbb{R}^n \times \mathbb{R}^{\hat{m}}) \times (\mathbb{R}^l \times \mathbb{R}^{\hat{m}}) \rightarrow \mathbb{R}$,

$$\hat{L}((x, s), (\lambda, v)) = f(x) + \langle (\lambda, v), \hat{h}(x, s) \rangle, \quad (2.11)$$

есть функция Лагранжа задачи (2.9). А именно, для текущего приближения $((x^k, s^k), (\lambda^k, v^k)) \in (\mathbb{R}^n \times \mathbb{R}^{\hat{m}}) \times (\mathbb{R}^l \times \mathbb{R}^{\hat{m}})$ к решению системы (2.10) следующее приближение $((x^{k+1}, s^{k+1}), (\lambda^{k+1}, v^{k+1}))$ ищется как решение линейной системы

$$\begin{pmatrix} \frac{\partial^2 \hat{L}}{\partial(x, s)^2}((x^k, s^k), (\lambda^k, v^k)) & (\hat{h}'(x^k, s^k))^T \\ \hat{h}'(x^k, s^k) & 0 \end{pmatrix} \begin{pmatrix} (x, s) - (x^k, s^k) \\ (\lambda, v) - (\lambda^k, v^k) \end{pmatrix} = - \begin{pmatrix} \frac{\partial \hat{L}}{\partial(x, s)}((x^k, s^k), (\lambda^k, v^k)) \\ \hat{h}(x^k, s^k) \end{pmatrix}. \quad (2.12)$$

В роли начального приближения для МНЛ используется точка $((\hat{x}, s^0), (\hat{\lambda}, v^0))$, где $s_i^0 = \sqrt{2 \max\{0, -g_i(\hat{x})\}}$, $v_i^0 = \hat{\mu}_i$, $i \in \hat{A}$. Для каждого $k = 1, \dots, K$ проверяется выполнение условия

$$\hat{\sigma}((x^k, s^k), (\lambda^k, v^k)) \leq \varepsilon, \quad (2.13)$$

где функция $\hat{\sigma} : (\mathbb{R}^n \times \mathbb{R}^m) \times (\mathbb{R}^l \times \mathbb{R}^m) \rightarrow \mathbb{R}_+$ представляет собой невязку системы Лагранжа (2.10) и определяется в виде

$$\hat{\sigma}((x, s), (\lambda, v)) = \left\| \begin{pmatrix} \frac{\partial \hat{L}}{\partial(x, s)}((x, s), (\lambda, v)), \hat{h}(x, s) \end{pmatrix} \right\|. \quad (2.14)$$

Если условие (2.13) выполнено для некоторого k , то далее проверяется выполнение условия (2.6) для $\tilde{x} = x^k$, $\tilde{\lambda} = \lambda^k$, $\tilde{\mu}_i = \max\{0, v_i^k\}$, $i \in \hat{A}$, $\tilde{\mu}_i = 0$, $i \in \{1, \dots, m\} \setminus \hat{A}$. Если это условие выполняется, то алгоритм завершает работу, поскольку найдено приближение $(\tilde{x}, \tilde{\lambda}, \tilde{\mu})$ нужного качества. В противном случае, если $k < K$, то ищется следующее приближение согласно (2.12); если же $k = K$, то запуск ускорителя объявляется неудачным.

В ALGENCAN по умолчанию используются $K = 10$, $\varepsilon = 10^{-8}$ и начальное значение $\hat{\varepsilon} = 10^{-4}$. Возможное правило уменьшения $\hat{\varepsilon}$ см. в [10]. Кроме того, ускоритель в ALGENCAN всегда запускается для исходной задачи, даже если алгоритм 1 запускается для масштабированного варианта (2.3). В последнем случае двойственное начальное приближение ускорителя $(\hat{\lambda}, \hat{\mu})$ модифицируется соответствующим образом.

Имея в виду возможное улучшение поведения описанной схемы применительно к задачам с нерегулярными ограничениями, рассмотрим также вариант ускорителя, в котором вместо итераций МНЛ для задачи (2.9) осуществляются итерации стабилизированного метода Ньютона–Лагранжа (СМНЛ), т.е. (2.12) заменяется линейной системой

$$\begin{pmatrix} \frac{\partial^2 \hat{L}}{\partial(x, s)^2}((x^k, s^k), (\lambda^k, v^k)) & (\hat{h}'(x^k, s^k))^T \\ \hat{h}'(x^k, s^k) & -\sigma_k I \end{pmatrix} \begin{pmatrix} (x, s) - (x^k, s^k) \\ (\lambda, v) - (\lambda^k, v^k) \end{pmatrix} = \\ = - \begin{pmatrix} \frac{\partial \hat{L}}{\partial(x, s)}((x^k, s^k), (\lambda^k, v^k)) \\ \hat{h}(x^k, s^k) \end{pmatrix}, \quad (2.15)$$

где $\sigma_k \geq 0$ – параметр стабилизации, выбор которого основан на оценке расстояния от $((x^k, s^k), (\lambda^k, v^k))$ до множества решений системы Лагранжа (2.10) задачи (2.9). А именно, будем полагать, что $\sigma_k = \hat{\sigma}((x^k, s^k), (\lambda^k, v^k))$.

Для СМНЛ, применяемого к задаче (2.9), справедлива следующая теорема о локальной сходимости, вытекающая из результатов в [1]. Эта теорема показывает, что если СМНЛ запускается в качестве ускорителя алгоритма 1 вблизи стационарной точки при двойственном приближении, достаточно близком к некритическому множителю Лагранжа, удовлетворяющему условию строгой дополненности, то имеет место сверхлинейная сходимость итераций ускорителя.

Теорема 2. Пусть функция $f: \mathbb{R}^n \rightarrow \mathbb{R}$ и отображения $h: \mathbb{R}^n \rightarrow \mathbb{R}^l$ и $g: \mathbb{R}^n \rightarrow \mathbb{R}^m$ дважды дифференцируемы в окрестности точки $\bar{x} \in \mathbb{R}^n$, и их вторые производные непрерывны в этой точке. Пусть $\bar{x}$ стационарная точка задачи (1.1), а $(\bar{\lambda}, \bar{\mu}) \in \mathcal{M}(\bar{x})$ – некритический множитель Лагранжа, удовлетворяющий условию строгой дополненности $\bar{\mu}_{A(\bar{x})} > 0$.

Тогда для любого $\hat{\varepsilon} > 0$ и любого начального приближения $(x^0, \lambda^0, \mu^0) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$, достаточно близкого к $(\bar{x}, \bar{\lambda}, \bar{\mu})$, если $\hat{A}$ определяется согласно (2.7) при $\hat{x} = x^0$, и если $s_i^0 = \sqrt{2 \max\{0, -g_i(x^0)\}}$, $v_i^0 = \mu_i^0$, $i \in \hat{A}$, то существует единственная траектория $\{((x^k, s^k), (\lambda^k, v^k))\}$ такая, что для каж-

до $k = 0, 1, \dots$ точка $\{(x^{k+1}, s^{k+1}), (\lambda^{k+1}, v^{k+1})\}$ удовлетворяет системе (2.15), в которой $\hat{h}$ определяется согласно (2.8), $\hat{L}$ — согласно (2.11), а $\sigma_k = \hat{\sigma}((x^k, s^k), (\lambda^k, v^k))$ — согласно (2.14); траектория сверхлинейно сходится к $((\bar{x}, \bar{s}), (\lambda^*, v^*))$, где $\bar{s}_i = \sqrt{-2g_i(\bar{x})}$, $v_i^* = \mu_i^*$, $i \in \hat{A}$, при некотором $(\lambda^*, \mu^*) \in \mathcal{M}(\bar{x})$, причем $\mu_i^* = 0$, $i \in \{1, \dots, m\} \setminus \hat{A}$.

Доказательство. Достаточная близость x^0 к $\bar{x}$ гарантирует включение $A(\bar{x}) \subset \hat{A}$. Отсюда следует, что $\bar{x}$ является стационарной точкой задачи

$$f(x) \rightarrow \min, \quad h(x) = 0, \quad g_i(x) \leq 0, \quad (2.16)$$

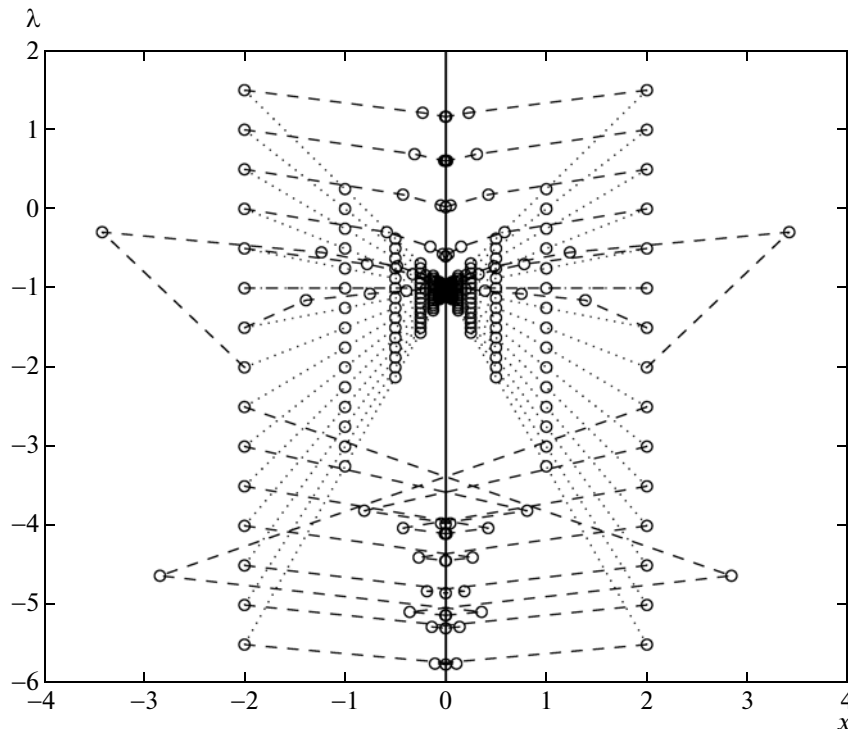
а $(\bar{\lambda}, \bar{\mu}_{\hat{A}})$ — отвечающий этой стационарной точке не критический множитель Лагранжа, для которого выполнено условие строгой дополнителности. Применяя к задаче (2.16) теорему 2 из [1], получаем требуемое утверждение.

Отметим, что при нарушении условия строгой дополнителности на сверхлинейную сходимость СМНЛ, применяемого к задаче (2.9), рассчитывать не приходится. Действительно, в этом случае к задаче (2.16) можно применить предложение 5 из [1], откуда следует, что в задаче (2.9) множитель $(\bar{\lambda}, \bar{\mu}_{\hat{A}})$, отвечающий стационарной точке $(\bar{x}, \bar{s})$, является критическим.

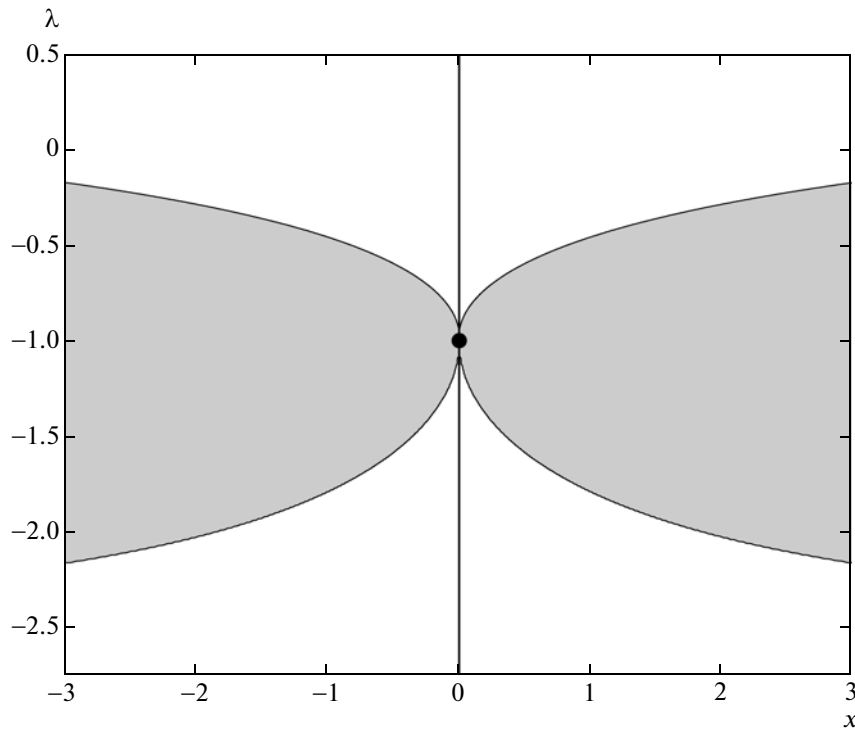
Сравнительное поведение МНЛ и СМНЛ при наличии критического множителя Лагранжа наглядно демонстрирует следующий простой

Пример 1. (Тест 20101 из коллекции DEGEN [22].) Пусть $n = l = 1$, $m = 0$, $f(x) = x^2$, $h(x) = x^2$. В задаче (1.1) с такими данными имеется единственная допустимая точка $\bar{x} = 0$, которая, соответственно, является единственным решением, причем $\mathcal{M}(\bar{x}) = \mathbb{R}$, а единственным критическим множителем является $\bar{\lambda} = -1$.

На фиг. 1 вертикальная сплошная линия соответствует множеству $\{\bar{x}\} \times \mathcal{M}(\bar{x})$, пунктирными линиями показаны траектории МНЛ, а штриховыми — траектории СМНЛ. Запуски выполнялись из точек (x^0, λ^0) , где x^0 равнялось -2 или 2 , а λ^0 принимало различные значения из сегмента



Фиг. 1.



Фиг. 2.

$[-5.5, 0.5]$. Траектории МНЛ всегда сходятся к $(\bar{x}, \bar{\lambda})$ с линейной скоростью. Для СМНЛ поведение двойственной траектории зависит от двойственной начальной точки: если $\lambda^0 \in (-2, -0.4)$ (интервал указан приблизительно), то наблюдается сходимость к $(\bar{x}, \bar{\lambda})$ с линейной скоростью; для остальных использованных значений λ^0 наблюдается сверхлинейная сходимость к $(\bar{x}, \lambda^*)$ при некотором $\lambda^* \neq \bar{\lambda}$.

На фиг. 2 серым цветом показана область начальных приближений (x^0, λ^0) , из которых СМНЛ сходится к критическому множителю Лагранжа.

3. ВЫЧИСЛИТЕЛЬНЫЙ ЭКСПЕРИМЕНТ

Приведены результаты вычислительного эксперимента, в котором исследовалось влияние критических множителей на скорость сходимости солвера ALGENCAN с различными способами управления параметром штрафа, а также на поведение двух описанных в предыдущем разделе ускорителей этого солвера: МНЛ и СМНЛ.

В эксперименте использовалась версия 2.3.7 пакета ALGENCAN с AMPL-интерфейсом, скомпилированная с использованием библиотеки MA57 (см. [23]). Данная библиотека предназначена для эффективного решения разреженных систем линейных уравнений, что позволяет существенно повысить общую эффективность солвера на задачах большой размерности.

Эксперимент проводился на двух тестовых коллекциях: DEGEN (см. [22]) и MacMPEC (см. [24]) (обе существуют в виде AMPL-моделей).

Коллекция DEGEN содержит 109 небольших задач математического программирования с нерегулярными ограничениями, из которых в эксперименте использовались 98. Остальные 11 задач были исключены по следующим причинам: задачи 20205, 20223, 30204 и 40207 имеют неограниченные на допустимом множестве целевые функции (хотя в них существуют стационарные точки, в которых ограничения нерегулярны); задачи 2DD01-50h, 2DD01-50v, 2DD01-500h, и 2DD01-500v слишком большие по сравнению с остальными; задачи 20201, 40210 и 40211 содержат только простые ограничения, вследствие чего ALGENCAN не делает для этих задач ни одной внешней итерации. Отметим, что в 52 из оставшихся 98 задач существуют критические относи-

тельно некоторого множества индексов $A \subset A(\bar{x})$ множители Лагранжа; в 33 задачах критических множителей нет; в 13 задачах решению вообще не отвечает ни один множитель Лагранжа, что вполне возможно в случае нерегулярных ограничений.

МасМРЕС представляет собой коллекцию из 180 задач оптимизации с комплементарными ограничениями, т.е. задач вида

$$f(x) \longrightarrow \min, \quad h(x) = 0, \quad g(x) \leq 0, \quad G(x) \geq 0, \quad H(x) \geq 0, \quad \langle G(x), H(x) \rangle \leq 0, \quad (3.1)$$

где комплементарные ограничения задаются с помощью отображений $G, H: \mathbb{R}^n \longrightarrow \mathbb{R}^s$. Разумеется, задача (3.1) — частный случай задачи математического программирования (1.1). Вместе с тем, как известно (см., например, [25, разд. 4.3]), наличие комплементарных ограничений неизбежно приводит к нерегулярности. МасМРЕС содержит как небольшие академические примеры, так и задачи прикладного происхождения, в том числе большой размерности.

В эксперименте использовалась 161 задача из МасМРЕС: 13 задач были исключены, поскольку в них присутствуют так называемые смешанные комплементарные ограничения; одна задача содержит двоичную переменную; остальные 5 исключенных задач не имеют допустимых точек.

3.1. Эксперимент с правилами для параметра штрафа: критические множители и скорость сходимости

Начнем с эксперимента, в котором сравнивались ALGENCAN как он есть (см. алгоритм 1 и сопровождающие его комментарии; в дальнейшем этот базовый вариант будем обозначать через M1), а также его модификация (в дальнейшем обозначаемая M2), в которой параметр штрафа увеличивался в два раза на каждой итерации, причем на первой итерации он брался равным 1. Таким образом, согласно теореме 1, для M2 можно ожидать сверхлинейную скорость сходимости.

В обоих вариантах M1 и M2 алгоритма использовались следующие настройки. Во-первых, был отключен ускоритель, поскольку в данном эксперименте изучается поведение “чистого” метода множителей. Во-вторых, ограничение на число внутренних итераций (т.е. итераций метода, применяемого для решения подзадач (2.1)) на первой внешней итерации было увеличено до 100 (значение по умолчанию 10). Кроме того, в экспериментах на коллекции DEGEN было отключено масштабирование, поскольку оно приводит к изменению множества множителей Лагранжа, отвечающих решению.

Для каждой задачи из коллекции DEGEN осуществлялось по 10 запусков каждого из алгоритмов M1 и M2 из случайных начальных точек. Для каждого из запусков анализировалась скорость убывания расстояния до прямого решения, а также притяжение двойственной траектории к критическим множителям (решения и критические множители указаны в моделях коллекции DEGEN).

На коллекции DEGEN результаты для M1 и M2 оказались практически одинаковыми. Это связано с тем, что для многих задач из DEGEN алгоритм 1, в котором подзадачи решаются точно (т.е. $\varepsilon_k = 0$ для всех k), находит точное решение за конечное число шагов. Для таких задач поведение алгоритма в значительной степени зависит от точности решения подзадач и мало зависит от правила изменения параметра штрафа. Поэтому ниже приводятся результаты только для M2.

Из 98 задач, использовавшихся в эксперименте, только для 80 наблюдалась сходимость к решениям, в которых нарушаются условия регулярности ограничений, и ниже обсуждаются результаты именно для этих задач. Для остальных задач либо запуски завершались неудачей, либо имела место сходимость к невырожденному решению.

Для всех задач, в которых множество множителей Лагранжа, отвечающих решению, пусто (таких задач 12 из 80), наблюдалась линейная сходимость прямой траектории, а норма двойственных приближений стремилась к бесконечности. Для тех задач, в которых решению отвечают только некритические множители (27 задач из 80), практически во всех случаях наблюдалась сверхлинейная сходимость прямой траектории, причем двойственная траектория сходилась к некоторому множителю Лагранжа.

Результаты для задач, в которых решению отвечает хотя бы один критический множитель (41 задача из 80), приведены в табл. 1. В столбце СЛ отмечены случаи сверхлинейной прямой сходимости, а в столбце КМ — случаи сходимости к критическому множителю. При этом черный кружок означает, что соответствующее поведение наблюдалось во всех запусках, а белый — что только для части запусков. Как видно из таблицы, сходимость к критическим множителям является типичной, но при этом сверхлинейная скорость прямой сходимости обычно сохраняется.

Таблица 1. Результаты для задач из DEGEN с критическими множителями

Задача	СЛ	КМ	Задача	СЛ	КМ
10201	●	●	20219	●	●
10203	●	●	20226	●	●
10206	●	●	20227	●	●
10207	●	●	20302	●	●
10301		●	20303	●	
20101	●		20304	●	
20103	●	●	20307		●
20105	○	○	20309		●
20107	●		2DD01-5v	●	
20109	●	●	30210	●	
20202	●	●	30211	●	
20203		●	30301	●	
20204	●		30401	●	●
20207	●	●	40201	●	
20209	○	○	40202	●	
20210	●	●	40205	●	●
20211	●		40206	●	●
20213	○	●	40208	●	●
20214		●	40401	●	
20216	●	●	40402	●	●
20217	○	●			

Вместе с тем, отсутствие сверхлинейной сходимости всегда сопряжено именно со сходимостью к критическим множителям.

Для задач из MacMPES осуществлялось по одному запуску каждого варианта алгоритма из указанных в моделях коллекции начальных точек. Для большинства этих задач ни решение, ни тем более отвечающие ему множители Лагранжа не известны. В связи с этим, под скоростью сходимости здесь понималась скорость убывания выдаваемой ALGENCAN на каждой итерации невязки системы ККТ.

Для варианта M2 были получены следующие результаты. Для 28 задач из 161 запуск завершился неудачей. Из оставшихся 133 задач лишь для 16 наблюдалась линейная скорость сходимости, а для остальных — сверхлинейная.

Базовый вариант M1 оказался значительно более робастным: лишь для 10 задач его запуск был неудачным. Однако, как и следовало ожидать, отсутствие сверхлинейной скорости сходимости наблюдалось значительно чаще, чем для M1, а именно, для 98 запусков из 151 успешных. Правда, при этом для 30 из этих задач линейная скорость была весьма быстрой: невязка системы ККТ убывала на каждой итерации более чем в 10 раз. Сверхлинейная скорость сходимости наблюдалась для 53 задач.

Из полученных результатов можно сделать следующие выводы. Если параметр штрафа стремится к бесконечности, то метод модифицированных функций во многих случаях сохраняет сверхлинейную скорость сходимости несмотря на нерегулярность ограничений решаемой задачи, и несмотря на притяжение двойственных траекторий к критическим множителям Лагранжа (при их наличии). Однако, как и следовало ожидать, бесконечное увеличение параметра штрафа негативно влияет на робастность метода.

3.2. Эксперимент с правилами для параметра штрафа: общая эффективность

Как отмечено выше, высокая скорость сходимости внешних итераций метода множителей может не компенсировать расходов на ее обеспечение: быстрый рост стоимости решения подзадач может оказывать решающее негативное влияние на общую эффективность (и даже на робастность) метода.

Для сравнения эффективности методов множителей с разными способами управления параметром штрафа варианты М1 и М2 алгоритма 1, описанные в предыдущем разделе, были дополнены следующими двумя. Вариант М3 аналогичен М2, но параметр штрафа увеличивается на каждой итерации в $\rho = 10$ раз. В варианте М4 параметр штрафа также увеличивается в 10 раз на каждой итерации, но, как и в базовом ALGENCAN, использовались рестарты для этого параметра, а значение c_0 определялось с помощью процедуры инициализации (см. разд. 2).

Похожие эксперименты проводились в [20], [21], где авторы исследовали влияние различных способов управления параметром штрафа на робастность и эффективность ALGENCAN. В [20] сравнивались базовый ALGENCAN и его модификации, в которых параметр штрафа растет на каждой итерации. В [21] исследовалось влияние на базовый ALGENCAN различных значений параметра ρ и рестартов. Однако, в отличие от настоящей работы, тестирование в [20], [21] проводилось на задачах с регулярными ограничениями.

В данном эксперименте по-прежнему был отключен ускоритель для всех вариантов алгоритма. Остальные настройки ALGENCAN были установлены по умолчанию. В частности, использовалось масштабирование, а ограничение на число внутренних итераций для первой внешней итерации было оставлено равным 10.

Сравнение вариантов ALGENCAN проводилось по следующим четырем показателям: количество внешних и внутренних итераций, а также количество вычислений целевой функции и ограничений. Количество внешних итераций отражает скорость сходимости алгоритма, а остальные характеристики могут рассматриваться как различные показатели его эффективности.

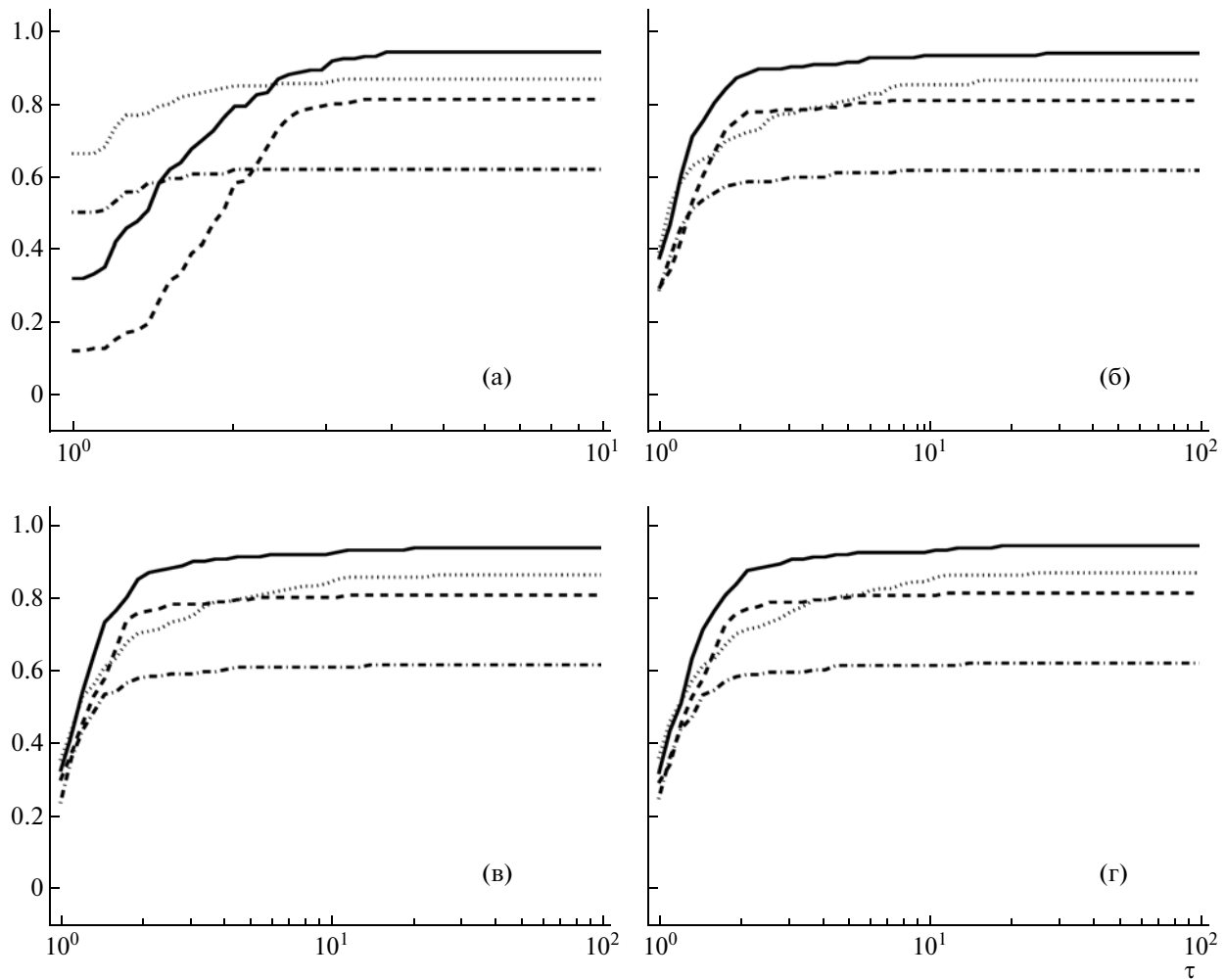
Результаты на коллекции MacMPEC представлены на фиг. 3 в форме так называемых “performance profiles” (см. [26]). Для каждого из алгоритмов значение функции, график которой изображен на рисунке, в точке $\tau \in [1, \infty)$ есть доля задач в наборе, на которых результат данного алгоритма (т.е. количество итераций или вычислений функций) был не более чем в τ раз хуже наилучшего результата для данной задачи среди всех сравниваемых алгоритмов. При этом считается, что результат неудачного запуска в бесконечное число раз хуже любого результата. В частности, значение этой функции при $\tau = 1$ соответствует доле запусков, на которых результат данного алгоритма был наилучшим. Значение функции при больших τ характеризует робастность алгоритма, т.е. долю его успешных запусков. Варианту М1 отвечает сплошная линия, М2 — штриховая, М3 — штрих-пунктирная, М4 — пунктирная.

Как видно из фиг. 3, для всех вариантов алгоритма, в которых параметр штрафа увеличивается на каждой итерации, робастность ниже, чем для базового ALGENCAN: в вариантах М2–М4 параметр штрафа чаще достигает величины 10^{20} , и в этих случаях запуск объявляется неудачным. При этом робастность варианта М4 выше, чем М2 и (особенно) М3, т.е. адаптивная инициализация параметра штрафа и его “рестарты” повышают вероятность успешности запуска.

Из фиг. 3а также следует, что варианты М3 и (особенно) М4 имеют лучший результат по количеству внешних итераций по сравнению с базовым вариантом М1, что свидетельствует о более высокой скорости сходимости М3 и (особенно) М4. Результат варианта М2, напротив, хуже, чем результат М1: хотя параметр штрафа в М2 растет на каждой итерации, это происходит слишком медленно.

С другой стороны, из фиг. 3б–3г следует, что все варианты алгоритма, в которых параметр штрафа увеличивается на каждой итерации (и даже те, которые демонстрируют более высокую скорость сходимости), не обыгрывают базовый ALGENCAN по эффективности. Отметим также, что вариант М4 несколько более эффективен, чем М2 и М3, хотя и незначительно.

Для представления результатов для коллекции DEGEN использовалась модификация подхода из [26], учитывающая многократность запусков алгоритмов для задач из этой коллекции. А именно, для каждого из алгоритмов на рисунках представлен график функции $\pi : [1, +\infty) \rightarrow [0, 1]$, определяемой следующим образом. Обозначим через k_p средний результат на один успеш-



Фиг. 3. а – внешние итерации, б – внутренние итерации, в – вычисления целевой функции, г – вычисления ограничений.

ный запуск, а через s_p долю успешных запусков данного алгоритма для p -й задачи в коллекции. Пусть r_p – наименьшее значение k_p по всем алгоритмам. Тогда для каждого $\tau \in [1, +\infty)$ полагаем

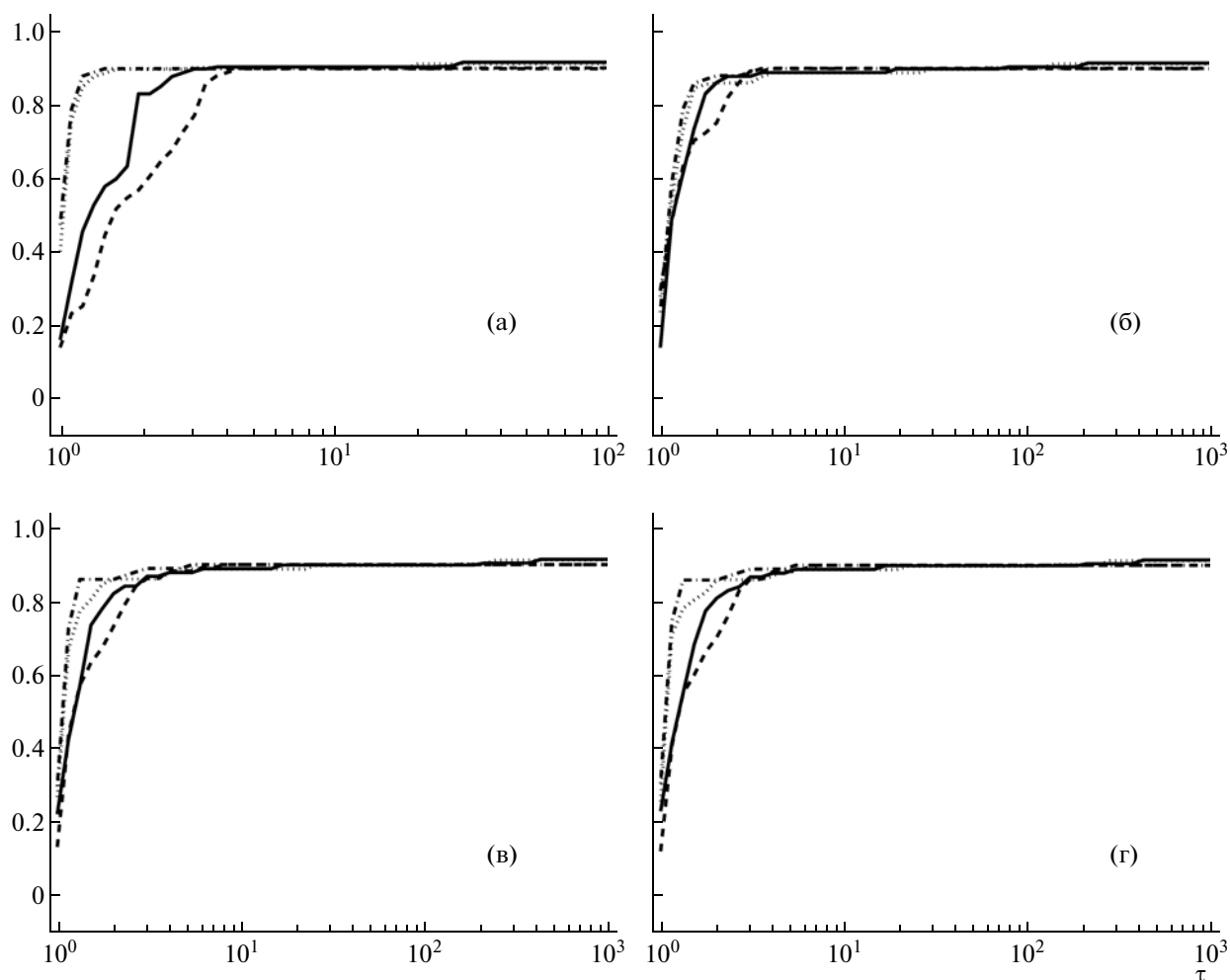
$$\pi(\tau) = \frac{1}{P} \sum_{p \in R(\tau)} s_p,$$

где P – общее количество задач в коллекции, а $R(\tau)$ – множество задач, для которых результат данного алгоритма не более чем в τ раз хуже наилучшего:

$$R(\tau) = \{p = 1, \dots, P \mid k_p \leq \tau r_p\}.$$

Из фиг. 4 (линии те же, что и на фиг. 3) видно, что на коллекции DEGEN все варианты алгоритма имеют примерно одинаковую робастность. Результаты по числу внешних итераций, показанные на фиг. 4а, повторяют картину для МасМРЕС, хотя различие между вариантами М1 и М2, и между М3 и М4 гораздо менее значительно. Однако, как следует из фиг. 4б–4г, в отличие от результатов для МасМРЕС, варианты М3 и М4 обгоняют М1 и М2 также и по остальным характеристикам, хотя и незначительно.

Приведенные результаты подтверждают высказанные выше опасения: высокая скорость сходимости метода множителей, достигаемая даже на задачах с нерегулярными ограничениями за счет бесконечного увеличения параметра штрафа, во многих случаях не приводит к высокой итоговой эффективности метода.



Фиг. 4. а — внешние итерации, б — внутренние итерации, в — вычисления целевой функции, г — вычисления ограничений.

3.3. Эксперимент с ускорителями

В заключение приведем результаты сравнения двух вариантов ALGENCAN с включенным ускорителем: базового варианта, в котором ускорителем является МНЛ, и варианта, использующего в качестве ускорителя СМНЛ. В данном эксперименте использовались настройки ALGENCAN по умолчанию.

Тестирование на коллекции DEGEN проводилось так же, как и в п. 3.1. Для 10 задач из 98 все запуски обоих тестируемых вариантов алгоритма заканчивались либо неудачей, либо сходимостью к решению, в котором ограничения регулярны. Еще для 14 ни один из ускорителей вообще не запускался ни разу: ALGENCAN успевал набрать требуемую точность (т.е. найти приближение, удовлетворяющее условию (2.6)) шагами метода множителей до того, как выполнялся тест (2.5) на включение ускорителя. Поэтому результаты приводятся для оставшихся 74 задач. Для некоторых из этих задач ускоритель запускался на нескольких внешних итерациях (т.е. на всех таких итерациях, кроме, возможно, последней, его запуски завершались неудачей). В таких случаях анализировалась только траектория, сгенерированная ускорителем на последней внешней итерации.

Для 11 из 12 задач, в которых решению не отвечает ни один множитель Лагранжа, запуски ускорителей всегда завершались неудачей. В оставшемся примере такого рода 2DD01-5h все запуски обоих ускорителей завершились успехом, и при этом во всех случаях скорость прямой сходимости была сверхлинейной.

Результаты для задач, в которых есть критические множители (38 задач из 74), представлены в табл. 2 (для МНЛ, и через косую черту для СМНЛ). В столбце УУ отмечены случаи успешных

Таблица 2. Результаты для задач из DEGEN с критическими множителями

Задача	УУ	СЛ	КМ	Задача	УУ	СЛ	КМ
10201	●/●	/●	●/	20214	●/●	/	●/●
10203	●/●	/	●/●	20216	●/●	/○	●/○
10206	●/●	/●	●/	20217	●/●	/	●/●
10207	●/●	/●	●/	20219	●/●	/	●/●
10301	/	/	/	20226	●/●	/○	●/●
10403	○/○	●/●	/	20227	●/●	/	●/●
20101	●/●	●/●	/	20302	●/●	/○	●/○
20103	●/●	/○	●/○	20303	●/●	○/●	○/
20105	●/●	●/●	/	20307	●/	/	●/
20107	●/●	○/●	○/	20309	●/●	/	●/●
20109	●/●	/	●/●	20312	○/○	/	●/●
20202	●/●	/	●/●	2DD01-5v	●/●	○/○	○/○
20203	○/○	/	●/●	30211	○/○	●/●	/
20204	●/●	○/●	○/	30301	●/●	●/●	○/
20207	●/●	/	●/●	30401	●/●	○/●	○/
20209	○/○	○/○	○/○	40201	○/○	●/●	/
20210	●/●	/	●/●	40206	/	/	/
20211	●/●	●/●	/	40208	/	/	/
20213	/●	/	/●	40402	●/●	/●	●/

запусков ускорителей, т.е. те случаи, когда решение находилось именно шагами ускорителя. В столбце СЛ отмечены случаи сверхлинейной прямой сходимости, а в столбце КМ — случаи притяжения двойственной траектории к критическим множителям. Черные и белые кружки имеют тот же смысл, что и в табл. 1.

Как видно из таблицы, для обоих ускорителей практически все запуски завершились успехом. Однако скорость сходимости и поведение двойственной траектории ускорителей несколько различаются.

Для МНЛ практически во всех случаях имела место сходимость к критическим множителям; при этом было лишь несколько случаев сверхлинейной прямой сходимости, причем практически во всех из них двойственная траектория сходилась к некритическому множителю. Для СМНЛ сверхлинейная сходимость прямой траектории наблюдалась значительно чаще, а двойственная траектория реже сходилась к критическому множителю, хотя таких случаев, как и случаев линейной прямой сходимости, по-прежнему большинство. Как и для МНЛ, практически во всех случаях сверхлинейной сходимости двойственная траектория сходилась к некритическому множителю. Таким образом, сходимость к критическим множителям в большинстве случаев разрушает сверхлинейную скорость сходимости обоих тестируемых ускорителей.

Следует отметить, что при наличии ограничений-неравенств аналогичный эффект может наблюдаться даже в том случае, когда решению отвечают только некритические множители, поскольку критические множители могут присутствовать в соответствующей задаче (2.9) со вспомогательными переменными, к которой применяется МНЛ или СМНЛ. Более того, несложно убедиться, что наличие критических множителей в задаче (2.9) является в определенном смысле типичным. Действительно, как следует из [1, предложение 5], если $A(\bar{x}) \subset \hat{A}$, то любому множителю Лагранжа $(\lambda, \mu) \in \mathcal{M}(\bar{x})$, для которого нарушается условие строгой дополнителности (т.е. $\mu_i = 0$ при некотором $i \in A(\bar{x})$), соответствует критический множитель Лагранжа $(\lambda, \mu_{\bar{A}})$ в задаче (2.9). Далее, из (1.2) очевидно, что множество

$$\mathcal{M}_A(\bar{x}) = \{ \mu_{A(\bar{x})} | (\lambda, \mu) \in \mathcal{M}(\bar{x}) \}$$

Таблица 3. Результаты для задач из DEGEN без критических множителей

Задача	УУ	СЛ	Задача	УУ	СЛ
00301	●/●	●/●	20222	○/●	●/●
00302	○/○	/	20225	○/○	○/○
00501	/	/	30202	●/○	●/●
10202	●/●	●/●	30205	●/●	●/●
10204	/●	/●	30206	○/○	●/●
10205	○/○	/	30303	●/●	●/●
20102	●/●	●/●	40203	/	/
20104	○/●	/●	40204	●/●	●/●
20106	/	/	01001	●/●	●/●
20110	●/●	●/●	10401	○/○	○/○
20111	/	/	10404	○/○	○/○
20221	●/●	●/●	10701	○/○	○/○

представляет собой пересечение аффинного множества в $\mathbb{R}^{|A(\bar{x})|}$ с неотрицательным ортантом, и поэтому выполнение $\mu_{A(\bar{x})} > 0$ для всех $\mu_{A(\bar{x})} \in \mathcal{M}_A(\bar{x})$ возможно только в том случае, когда $\mathcal{M}_A(\bar{x})$ состоит из единственного элемента. Последнее означает, что вектор μ одинаковый для всех множителей $(\lambda, \mu) \in \mathcal{M}(\bar{x})$, что нетипично для задач с нерегулярными ограничениями.

Результаты для задач, в которых решению отвечают только не критические множители (24 задачи из 74), приведены в табл. 3. Отметим, что практически для всех из них в соответствующих задачах со вспомогательными переменными присутствуют критические множители. Вместе с тем, картина несколько отличается от представленной в табл. 2. Результаты для двух используемых ускорителей почти не различаются. При этом доля неудачных запусков ускорителей весьма существенна, однако успешных запусков все же значительно больше. Кроме того, для успешных запусков скорость прямой сходимости практически всегда была сверхлинейной.

Поведение ускорителей, демонстрируемое табл. 2 и 3, связано еще с одной эвристикой, реализованной в ALGENCAN, которая состоит в адаптивном возмущении диагональных элементов матриц, стоящих в левой части систем (2.12) и (2.15), с целью преодоления возможной вырожденности этих матриц (см. [10]). Возмущение диагональных элементов матрицы в (2.12) приближает поведение МНЛ к СМНЛ, и для некоторых задач приводит к сверхлинейной прямой сходимости МНЛ. Отключение данной эвристики приближает поведение ускорителей к ожидаемому: в частности, прямая сходимость МНЛ во многих случаях становится линейной, при этом двояственная траектория сходится к одному из критических множителей Лагранжа задачи (2.9). В некоторых случаях прямая сходимость МНЛ остается сверхлинейной, что связано с наличием вспомогательных переменных: прямая сходимость по всем переменным, включая вспомогательные, тем не менее, является линейной, как и сходимость невязки (2.14) системы Лагранжа (2.10).

Тестирование на MacMPEC проводилось так же, как в п. 3.1. Для МНЛ результаты следующие: для 22 задач из 161 не было сделано ни одного запуска ускорителя; еще для 29 задач все запуски ускорителя завершились неудачей; для 82 задач из оставшихся 110 имела место сверхлинейная сходимость невязки системы Лагранжа вдоль траектории ускорителя. Для СМНЛ ускоритель также не запускался ни разу для 22 задач; для 39 задач все запуски ускорителя были неудачными; из оставшихся 100 задач сверхлинейная сходимость также наблюдалась в большинстве случаев (для 74 задач). Отметим, что сверхлинейная сходимость МНЛ во многих случаях связана с описанной выше эвристикой, реализованной в ALGENCAN: ее отключение практически всегда приводит к линейной скорости сходимости.

Сравнение ALGENCAN с двумя вариантами ускорителя по робастности и эффективности, аналогичное эксперименту в п. 3.2, также не выявляет существенных различий в их поведении.

Таким образом, замена используемого в ALGENCAN стандартного ускорителя МНЛ на СМНЛ не приводит к сколько-нибудь существенному повышению эффективности или робастности солвера.

4. ЗАКЛЮЧЕНИЕ

Несмотря на теоретическую привлекательность методов множителей как стратегии решения задач оптимизации с потенциально нерегулярными ограничениями, высокие затраты на решение подзадач (особенно на поздних итерациях, когда требуемая точность решения этих подзадач высока, и при больших значениях параметра штрафа), не позволяют эффективно совместить для задач данного класса глобальную сходимость этих методов с высокой скоростью сходимости. Более того, простейшие модификации методов множителей, которые заключаются в использовании на финальной фазе различных ускорителей (таких как метод Ньютона—Лагранжа и его стабилизированная версия), также оказываются недостаточными для достижения этой цели. Основной причиной этого является притяжение двойственных траекторий как самих методов множителей, так и ускорителей к критическим множителям Лагранжа.

Таким образом, для получения действительно эффективного глобально сходящегося алгоритма для задач оптимизации с нерегулярными ограничениями недостаточно использовать стабилизированный метод Ньютона—Лагранжа лишь в качестве ускорителя метода множителей: требуются другие принципы интеграции этих методов, разработка которых является темой дальнейших исследований авторов.

СПИСОК ЛИТЕРАТУРЫ

1. *Izmailov A.F., Solodov M.V.* Stabilized SQP revisited // *Math. Program.* 2012. V. 122. № 1. P. 93–120.
2. *Измаилов А.Ф.* Об аналитической и вычислительной устойчивости критических множителей Лагранжа // *Ж. вычисл. матем. и матем. физ.* 2005. Т. 45. № 6. С. 966–982.
3. *Izmailov A.F., Solodov M.V.* Examples of dual behaviour of Newton-type methods on optimization problems with degenerate constraints // *Comput. Optimizat. Appl.* 2009. V. 42. № 2. P. 231–264.
4. *Hestenes M.R.* Multiplier and gradient methods // *J. Optimizat. Theory Appl.* 1969. V. 4. P. 303–320.
5. *Powell M.J.D.* A method for nonlinear constraints in minimization problems // *Optimization.* New York: Academic Press, 1969. P. 283–298.
6. *Бертсекас Д.* Условная оптимизация и методы множителей Лагранжа. М.: Радио и связь, 1987.
7. *Измаилов А.Ф., Солодов М.В.* Численные методы оптимизации. 2-е изд., перераб. и доп. М.: Физматлит, 2008.
8. LANCELOT. <http://www.cse.scitech.ac.uk/nag/lancelot/lancelot.shtml>.
9. ALGENCAN. <http://www.ime.usp.br/~egbirgin/tango/>.
10. *Birgin E.G., Martínez J.M.* Improving ultimate convergence of an augmented Lagrangian method // *Optimizat. Meth. Software.* 2008. V. 23. № 2. P. 177–195.
11. *Andreani R., Birgin E.G., Martínez J.M., Schuverdt M.L.* On augmented Lagrangian methods with general lower-level constraints // *SIAM J. Optimizat.* 2007. V. 18. № 4. P. 1286–1309.
12. *Andreani R., Haeser G., Schuverdt M.L., Silva P.J.S.* A relaxed constant positive linear dependence constraint qualification and applications // *Math. Program.* 2011. DOI: 10.1007/s10107-011-0456-0.
13. *Izmailov A.F., Solodov M.V., Uskov E.I.* Global convergence of augmented Lagrangian methods applied to optimization problems with degenerate constraints, including problems with complementarity constraints // *SIAM J. Optim.* Accepted for publication.
14. *Fernández D., Solodov M.V.* Local convergence of exact and inexact augmented Lagrangian methods under the second-order sufficient optimality condition // *SIAM J. Optim.* 2012. V. 22. P. 384–407.
15. *Izmailov A.F., Solodov M.V.* On attraction of Newton-type iterates to multipliers violating second-order sufficiency conditions // *Math. Program.* 2009. V. 117. № 1–2. P. 271–304.
16. *Izmailov A.F., Solodov M.V.* On attraction of linearly constrained Lagrangian methods and of stabilized and quasi-Newton SQP methods to critical multipliers // *Math. Program.* 2011. V. 128. № 2. P. 231–257.
17. *Измаилов А.Ф.* О предельных свойствах двойственных траекторий метода множителей Лагранжа // *Ж. вычисл. матем. и матем. физ.* 2011. Т. 51. № 1. С. 3–23.
18. *Wright S.J.* Superlinear convergence of a stabilized SQP method to a degenerate solution // *Comput. Optimizat. Appl.* 1998. V. 11. P. 253–275.
19. *Голишников М.М., Измаилов А.Ф.* Ньютоновские методы для задач условной оптимизации с нерегулярными ограничениями // *Ж. вычисл. матем. и матем. физ.* 2006. Т. 46. № 8. С. 1369–1391.
20. *Birgin E.G., Fernandez D., Martínez J.M.* On the boundedness of penalty parameters in an augmented Lagrangian method with lower level constraints // *Optimizat. Meth. Software.* 2012. V. 27. P. 1001–1024.
21. *Birgin E.G., Martínez J.M.* Augmented Lagrangian method with nonmonotone penalty parameters for constrained optimization // *Comput. Optimizat. Appl.* 2012. V. 51. № 3. P. 941–965.
22. DEGEN. http://w3.impa.br/~optim/DEGEN_collection.zip.
23. The HSL Mathematical Software Library. <http://www.hsl.rl.ac.uk>.
24. MacMPEC. <http://wiki.mcs.anl.gov/leyffer/index.php/MacMPEC>.
25. *Измаилов А.Ф.* Чувствительность в оптимизации. М.: Физматлит, 2006.
26. *Dolan E., Moré J.* Benchmarking optimization software with performance profiles // *Math. Program.* 2002. V. 91. № 2. P. 201–213.