УДК 519.8

Ю.П. ЛАПТИН, О.А. БЕРЕЗОВСКИЙ

ИСПОЛЬЗОВАНИЕ КОНИЧЕСКОЙ РЕГУЛЯРИЗАЦИИ ПРИ ВЫЧИСЛЕНИИ ЛАГРАНЖЕВЫХ ОЦЕНОК В ЗАДАЧАХ КВАДРАТИЧНОЙ ОПТИМИЗАЦИИ

Аннотация. Для невыпуклых задач квадратичной оптимизации рассматривается вычисление оценок значений глобальных экстремумов на основе лагранжевых релаксаций исходных задач. На границе допустимой области оценочной задачи ее функции являются разрывными и плохо обусловленными, что накладывает определенные требования на вычислительные алгоритмы. Для учета указанных особенностей разработан новый подход, основанный на использовании конических регуляризаций выпуклых задач оптимизации. Он позволяет построить эквивалентную задачу безусловной оптимизации, целевая функция которой определена на всем пространстве переменных задачи и удовлетворяет условию Липшица.

Ключевые слова: задачи квадратичной оптимизации, лагранжева релаксация, условие неотрицательной определенности матрицы, коническая регуляризация.

ВВЕДЕНИЕ

Настоящая работа посвящена памяти выдающегося ученого в области математической кибернетики, одного из основателей теории недифференцируемой оптимизации — академика НАН Украины Наума Зуселевича Шора. В этом году ему исполнилось бы 80 лет. Один из разделов его научного наследия связан с исследованием общего класса задач квадратичной оптимизации (максимизации квадратичной формы при квадратичных ограничениях). Для вычисления оценок значений глобальных экстремумов в таких задачах применяются лагранжевые релаксации (например, [1, 2]), а также SDP-релаксации (релаксации к задачам полуопределенного программирования) [3], которые в некотором смысле эквивалентны [4, 5]. При использовании лагранжевых релаксаций формируются оценочные задачи в пространстве двойственных переменных, допустимые множества которых находятся в области неположительной определенности линейно зависящих от них матриц функций Лагранжа. Во внутренних точках допустимых множеств функции оценочных задач непрерывно дифференцируемы и принимают конечные значения. На границе области отрицательной определенности эти функции разрывны (при наличии положительных собственных чисел принимают значения $+\infty$), а вблизи границы матрицы функции плохо обусловлены (количество близких к нулю собственных чисел может приближаться к размерности пространства). В данной статье предлагаются новые подходы учета этих особенностей, основанные на использовании конических регуляризаций выпуклых задач оптимизации [6, 7].

В разд. 1 описываются постановки задачи квадратичной оптимизации и ее лагранжевой релаксации. В разд. 2 дан обзор численных алгоритмов решения двойственной задачи. В разд. 3 рассматривается обобщение конических регуляризаций, изложенных в [6], для используемых оценочных задач, двойственных к исходной задаче квадратичной оптимизации. В разд. 4 приводятся новые результаты, которые позволяют реализовать эффективные алгоритмы одномерного поиска при использовании конических регуляризаций для рассматриваемых задач. Эти результаты идейно близки алгоритму «граничного оракула», введенному в [8]. Также рассматриваются эффективные процедуры вычисления функций и субградиентов регуляризованной оценочной задачи.

1. КВАДРАТИЧНЫЕ ОПТИМИЗАЦИОННЫЕ ЗАДАЧИ И ОЦЕНКИ ИХ ОПТИМАЛЬНЫХ ЗНАЧЕНИЙ

Постановка квадратичной оптимизационной задачи

$$K^* = \sup K_0(\mathbf{x}), \quad \mathbf{x} \in R^n, \tag{1}$$

при ограничениях

$$K_i(\mathbf{x}) \le 0, \ i = 1, \dots, m_1; \ K_j(\mathbf{x}) = 0, \ j = m_1 + 1, \dots, N, \ m_1 \le N, \tag{2}$$

где $K_i(\mathbf{x}) = \langle \mathbf{A}_i \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{l}_i, \mathbf{x} \rangle + c_i$, $\mathbf{A}_i$ — симметричные матрицы, $\mathbf{l}_i$ — векторы соответствующих размерностей, $c_i \in R$, $i \in \{0, 1, \dots, N\}$.

К виду задачи (1), (2) сводится ряд общеизвестных задач, к которым относятся задача линейной дополнительности, квадратичная задача о назначениях, задачи проектирования и упаковки, экстремальные задачи на графах (например, задачи раскраски и разбиения графов, нахождение максимального независимого множества, максимального k-клаба и т.п.). Кроме того, к квадратичным задачам сводится более широкий класс полиномиальных экстремальных задач (в которых целевая функция и все функции ограничений полиномиальные), решение систем полиномиальных уравнений с вещественными и булевыми переменными и т.д.

Задача (1), (2) является многоэкстремальной в общем случае. Для вычисления оценки оптимального значения используют лагранжевую релаксацию этой задачи по всем ограничениям. Отметим, что под лагранжевой релаксацией произвольной оптимизационной задачи

$$f^* = \inf_{\mathbf{x}} \{ f_0(\mathbf{x}) : f_i(\mathbf{x}) \le 0, i \in I^{LQ}, \ f_i(\mathbf{x}) = 0, i \in I^{EQ},$$

$$g_j(\mathbf{x}) \le 0, j \in J^{LQ}, \ g_j(\mathbf{x}) = 0, j \in J^{EQ} \}$$

в общем случае понимают задачу вида

$$f^* = \sup_{\mathbf{u} \in U^+} \inf_{\mathbf{x}} \left\{ f_0(\mathbf{x}) + \sum_{i \in I^{LQ}} u_i f_i(\mathbf{x}) + \sum_{i \in I^{EQ}} u_i f_i(\mathbf{x}) : \right.$$

$$\left. g_j(\mathbf{x}) \le 0, \ j \in J^{LQ} g_j(\mathbf{x}) = 0, j \in J^{EQ} \right\},$$

где $U^+ = \{\mathbf{u} : u_i \ge 0, \ i \in I^{LQ}, \ -\infty < u_i < +\infty, \ i \in I^{EQ}\}$.

Впервые термин «лагранжева релаксация» был предложен Джеоффрионом [9] применительно к задачам целочисленного программирования.

Запишем двойственную функцию для задачи (1), (2)

$$\psi(\mathbf{u}) = \sup_{\mathbf{x} \in R^n} L(\mathbf{x}, \mathbf{u}), \tag{3}$$

где

$$L(\mathbf{x}, \mathbf{u}) = K_0(\mathbf{x}) + \sum_{i=1}^{N} u_i K_i(\mathbf{x}) = \langle \mathbf{A}(\mathbf{u}) \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{l}(\mathbf{u}), \mathbf{x} \rangle + c(\mathbf{u}),$$

$$\mathbf{A}(\mathbf{u}) = \mathbf{A}_0 + \sum_{i=1}^{N} u_i \mathbf{A}_i, \ \mathbf{l}(\mathbf{u}) = \mathbf{l}_0 + \sum_{i=1}^{N} u_i \mathbf{l}_i, \ c(\mathbf{u}) = c_0 + \sum_{i=1}^{N} u_i c_i.$$

Пусть $U^- = \{\mathbf{u} = (u_1, \ldots, u_N): u_i \le 0, i = 1, \ldots, m_1\}$, T — множество допустимых решений задачи (1), (2), $\mathbf{u} \in U^-$, тогда

$$\psi(\mathbf{u}) = \sup_{\mathbf{x} \in R^n} L(\mathbf{x}, \mathbf{u}) \ge \sup_{\mathbf{x} \in T} L(\mathbf{x}, \mathbf{u}) \ge \sup_{\mathbf{x} \in T} K_0(\mathbf{x}) = K^*.$$

Положим D ($\overline{D}$) — подмножество R^N, состоящее из таких $\mathbf{u} \in U^-$, что $\mathbf{A}(\mathbf{u})$ является отрицательно-определенной (соответственно неположительно-определенной) матрицей,

$$D = \{\mathbf{u}: \lambda_{\max}(\mathbf{A}(\mathbf{u})) < 0,\ \mathbf{u} = (u_1, \ldots, u_N): u_i \le 0,\ i = 1, \ldots, m_1\},\ \overline{D} = \mathrm{cl} D.$$

Здесь $\lambda_{\max}(\mathbf{A}(\mathbf{u}))$ — максимальное собственное число матрицы $\mathbf{A}(\mathbf{u})$. Если $\mathbf{u} \notin \overline{D}$, то $\psi(\mathbf{u}) = +\infty$. В качестве оценки оптимального значения K^* используется величина [2]

$$\psi^* = \inf_{\mathbf{u} \in U^-} \psi(\mathbf{u}) = \inf_{\mathbf{u} \in \overline{D}} \psi(\mathbf{u}). \tag{4}$$

Задача (4) является задачей выпуклой оптимизации.

Пусть $\mathbf{u} \in D$, $\mathbf{x}(\mathbf{u})$ — решение задачи (3). Вектор $\mathbf{x}(\mathbf{u})$ является решением системы уравнений

$$2\mathbf{A}(\mathbf{u})\mathbf{x} + \mathbf{l}(\mathbf{u}) = 0. \tag{5}$$

Отсюда $\mathbf{x}(\mathbf{u}) = -\frac{1}{2}\mathbf{A}(\mathbf{u})^{-1}\mathbf{l}(\mathbf{u})$, $\psi(\mathbf{u}) = -\frac{1}{4}\langle \mathbf{A}(\mathbf{u})^{-1}\mathbf{l}(\mathbf{u}), \mathbf{l}(\mathbf{u})\rangle + c(\mathbf{u}) = \frac{1}{2}\langle \mathbf{x}(\mathbf{u}), \mathbf{l}(\mathbf{u})\rangle + c(\mathbf{u})$.

Во внутренних точках множества $\overline{D}$ функция $\psi(\mathbf{u})$ непрерывно дифференцируема и принимает конечные значения. Градиент $\mathbf{g}_\psi(\mathbf{u})$ функции $\psi(\mathbf{u})$ в точке $\mathbf{x} = \mathbf{x}(\mathbf{u})$ определяется соотношением [2]

$$\mathbf{g}_\psi(\mathbf{u}) = \mathbf{K}(\mathbf{x}(\mathbf{u})), \tag{6}$$

где $\mathbf{K}(\mathbf{x})$ — вектор, компонентами которого являются функции $K_i(\mathbf{x})$, $i = 1, \ldots, N$, формирующие ограничения задачи (1), (2).

На границе множества $\overline{D}$ функция $\psi(\mathbf{u})$ может принимать как конечные значения, так и значения $+\infty$. Особенности поведения функции $\psi(\mathbf{u})$ вблизи границы множества $\overline{D}$ накладывают определенные требования к вычислительным алгоритмам для решения задачи (4).

2. ЧИСЛЕННЫЕ АЛГОРИТМЫ РЕШЕНИЯ ДВОЙСТВЕННОЙ ЗАДАЧИ

Для упрощения изложения будем далее предполагать, что допустимая область двойственных переменных $\overline{D} = \{\mathbf{u}: \lambda_{\max}(\mathbf{A}(\mathbf{u})) \le 0,\ \mathbf{u} = (u_1, \ldots, u_N): u_i \le 0,$ $i = 1, \ldots, m_1\}$ в задаче (4) нахождения двойственной оценки содержит внутренние точки. При $\mathbf{u} \in D = \mathrm{int}\,(\overline{D})$ внутренняя задача (3) представляет собой задачу максимизации строго вогнутой квадратичной функции, решение которой $\mathbf{x}(\mathbf{u}) = \arg\max_{\mathbf{x} \in R^n} L(\mathbf{x}, \mathbf{u})$ однозначно определяется из невырожденной системы уравнений $2\mathbf{A}(\mathbf{u})\mathbf{x} + \mathbf{l}(\mathbf{u}) = 0$. Внешняя задача (4), как уже отмечалось, является задачей минимизации выпуклой двойственной функции $\psi(\mathbf{u})$ на выпуклых ограничениях двух типов:

а) $U^- = \{\mathbf{u} = (u_1, \ldots, u_N): u_i \le 0, i = 1, \ldots, m_1\}$ — отрицательность двойственных переменных, соответствующих ограничениям-неравенствам;

б) $\{\mathbf{u}:\mathbf{A}(\mathbf{u})\prec= 0\}=\{\mathbf{u}:\lambda_{\max}(\mathbf{A}(\mathbf{u}))\le 0\}$ — неположительная определенность матрицы $\mathbf{A}(\mathbf{u})$, что эквивалентно неположительности ее максимального собственного числа $\lambda_{\max}(\mathbf{A}(\mathbf{u}))\le 0$, функция которого является негладкой.

Таким образом, для решения задачи (4) можно использовать любой из численных методов решения задачи негладкой выпуклой оптимизации, например алгоритмы субградиентного типа или модификации метода секущих плоскостей.

Разные алгоритмы отличаются способами проверки и учета ограничений задачи. Например, проверку условия б) можно проводить с помощью правила Сильвестра [10], разложения Холецкого или другой процедуры разложения матрицы, соответствующей выделению полных квадратов [2, с. 93], нахождения максимального собственного числа [11, 12]. Среди способов учета можно назвать использование барьерных штрафных функций $\varepsilon_0/\lambda_{\max}(\mathbf{A}(\mathbf{u}))$ (см., например, [11]), где ε_0 — достаточно малое положительное число. Также можно на каждом шаге итерационного процесса при движении по направлению дробить шаг, чтобы всегда оставаться внутри области $\{\mathbf{u}:\mathbf{A}(\mathbf{u})\prec 0\}$ (см. [2, с. 97], этот способ был реализован при вычислительных экспериментах в [13]). В [12] применялся r-алгоритм с адаптивной регулировкой шага для минимизации штрафной функции

$$\Psi(\mathbf{u})=\begin{cases}\psi(\mathbf{u}), & \text{если } \mathbf{A}(\mathbf{u})\prec 0,\\ -\lambda_{\max}(\mathbf{A}(\mathbf{u})) & \text{в противном случае.}\end{cases}$$

Ограничения а) можно учитывать, применяя барьерные штрафные функции $\sum_{i=1}^{m_1}\varepsilon_i/u_i$ [11], где $\varepsilon_i,\ i\in\{1,\ldots,m_1\}$, — достаточно малые положительные числа, штрафные функции в форме функции максимума $-S_i\max\{0,u_i\}$, где S_i — достаточно большие положительные числа, а также используя в функции Лагранжа модуль двойственных переменных $L(\mathbf{x},-|\mathbf{u}|)$ [12] и т.д. Кроме того [14], ограничения этого типа можно исключить заменой условий в виде неравенств на равенства, добавляя в каждое из них квадрат вспомагательной переменной. Также в [14] описан подход, суть которого заключается в сведении квадратичной задачи общего вида (1), (2) к однородной квадратичной задаче, для которой решение внутренней задачи тривиально: $\mathbf{x}^*=\mathbf{0}$, где $\mathbf{0}$ — нулевой n-мерный вектор. В результате получаем формулировку двойственной SDP-релаксации исходной квадратичной задачи, к которой эффективно можно применять как методы полуопределенного программирования, так и методы негладкой оптимизации (например, в [14] задача сводится с помощью точной штрафной функции к безусловной задаче минимизации выпуклой недифференцируемой функции).

Далее приведем краткое описание общих методов оптимизации, которые могут использоваться при решении рассматриваемых задач. Чтобы не сузить возможности тех или иных представлений исходной задачи, будем использовать следующие обозначения: рассматривается общая задача негладкой оптимизации $f^*=\min_{\mathbf{x}\in T}f(\mathbf{x})$, $T\subseteq R^n$ — допустимое множество задачи, $\mathbf{g}(\mathbf{x})$ — произвольный субградиент целевой функции $f(\mathbf{x})$ в точке $\mathbf{x}$.

Субградиентный метод. В случае условной оптимизации итерационный процесс задается следующей схемой:

$$\mathbf{x}_{k+1}=\pi_T\left(\mathbf{x}_k-h_k\frac{\mathbf{g}(\mathbf{x}_k)}{\|\mathbf{g}(\mathbf{x}_k)\|}\right),\ h_k>0,\ h_k\to 0,\ \sum_{k=1}^{\infty}h_k=\infty,$$

где $\pi_T(\mathbf{x})$ — оператор проектирования на множество T. Отметим возможность использования вариантов, когда часть ограничений учитывается с помощью штрафов в целевой функции и соответственно проективное множество расширяется. В случае сведения исходной задачи к задаче безусловной негладкой оптимизации процедура проекции исключается.

Метод эллипсоидов. Суть метода состоит в следующем. Перед началом очередной итерации оптимум локализован в эллипсоиде с центом в точке $\mathbf{x}_k$. Если $\mathbf{x}_k$ является допустимой точкой, то с использованием субградиента целевой функции строится отсекающая гиперплоскость и вокруг оставшейся части эллипсоида описывается новый эллипсоид по возможности меньшего объема. Если же в точке $\mathbf{x}_k$ нарушено ограничение задачи, то для построения отсекающей гиперплоскости используется субградиент к этому ограничению.

Схемы этих двух классических методов являются достаточно жесткими. Теоретические оценки их скорости сходимости невысокие. На практике их скорость сходимости соответствует теоретическим оценкам даже тогда, когда применяются для минимизации очень простых функций. Далее рассмотрим алгоритмические схемы, которые являются более гибкими, чем субградиентный метод и метод эллипсоидов.

***r*-алгоритм.** Идея этого семейства алгоритмов состоит в шаге по антисубградиенту в пространстве, полученном из пространства на предыдущем шаге, которое растянуто по разности субградиентов, вычисленных в текущей и предыдущей точках. Наибольшее распространение получил вариант алгоритма с адаптивной регулировкой шага [2, 12, 15], как обладающий быстрой практической сходимостью и способностью практически всегда противостоять овражности.

Bundle-методы. Для алгоритмов этого класса характерно использование на текущей итерации информации, собранной на предыдущих шагах. Их множество достаточно многообразно, поэтому в качестве примера ограничимся рекуррентной формулой одного из них — метода Келли:

$$\mathbf{x}_{k+1} = \arg\min_{\mathbf{x}\in T}\left(\max_{1\le i\le k}\left(f(\mathbf{x}_i) + \langle \mathbf{g}(\mathbf{x}_k), \mathbf{x} - \mathbf{x}_i\rangle\right)\right).$$

Кроме рассмотренных выше алгоритмов, можно использовать особенности задачи (4), в частности известное свойство взаимосвязи двойственных оценок и SDP-релаксаций в квадратичных экстремальных задачах [5, 16]. Вместо вычисления оптимального значения двойственной функции (задача (4)) можно решать задачи полуопределенного программирования, как полученную путем SDP-релаксации, так и двойственную к ней, о которой уже упоминалось выше. В настоящее время существуют эффективные численные процедуры решения таких задач, использующие методы внутренней точки [16, 17], которые обладают как практической, так и теоретической эффективностью [17].

Методы внутренних точек. Принимая во внимание многоообразие алгоритмов данного класса, ограничимся наиболее общим определением из [18], согласно которому методы внутренних точек объединяют «семейство алгоритмов оптимизации, осуществляющих улучшение решения внутри области, состоящей из векторов, удовлетворяющих ограничениям-неравенствам в строгой форме».

В [19] рассматривается один из вариантов двойственного метода внутренней точки, который можно отнести к двойственным аффинно-масштабирующим методам. Описанный метод является обобщением на случай линейной задачи полуопределенного программирования двойственного барьерно-проективного метода, предложенного ранее для задач линейного программирования [20]. Его можно также рассматривать как двойственный аналог прямого метода из [21].

Главная цель [19] — показать, что перенесение соответствующего двойственного метода из [20] на задачи полуопределенного программирования не нарушает его сходимости, однако это происходит при достаточно жестких предположениях относительно решений прямой и двойственной задач. Для обоснования сходимости предлагаемого метода используется теорема Островского, в которой рассматривается линейное приближение алгоритмического оператора. Как известно, не всегда обобщения методов линейного программирования на линейные задачи полуопределенного программирования сохраняют сходимость [22].

Методы симплексного типа. В [23] предложено довольно универсальное обобщение симплекс-метода для задач с ограничениями, задаваемыми в форме линейных матричных неравенств. Для задач конического программирования с произвольными замкнутыми выпуклыми конусами обобщение симплекс-метода рассматривалось в [24]. В [25] предложен метод аффинно-масштабирующего типа, в котором допускался выход на границу допустимого множества и, следовательно, движение по его граням, а также переход с одной грани на другую.

В [26] рассмотрен вариант прямого симплекс-метода, аналогичный используемому в линейном программировании для перехода из одной крайней точки в другую. Основное внимание в этой работе уделяется переходу в нерегулярном случае, т.е. когда количество ограничений типа равенства может не совпадать с количеством переменных в крайней точке, равному так называемому «треугольному» числу (числу элементов симметричных матриц, расположенных на диагонали и под ней).

HR-алгоритм. В [8] предлагается итеративный метод решения SDP-задач, основанный на идеях, использующих понятие граничного оракула, случайные блуждания и оценку центра тяжести выпуклого множества допустимой области параметров. Этот подход распространяется и на другие типы задач выпуклой оптимизации.

Пусть $\mathbf{z}_0$ — начальная точка, лежащая внутри области D. Из текущей точки $\mathbf{z}_i$ направление движения выбирается случайно равномерно на единичной сфере, рассматривается одномерная прямая $\mathbf{z}_i + t\mathbf{r}_i$ и определяются точки ее пересечения с границей области D. Следующая HR-точка $\mathbf{z}_{i+1}$ выбирается случайно равномерно на полученной хорде. В [27, 28] доказано, что таким образом сформированная последовательность случайных векторов образует дискретную марковскую цепь, для которой выполняется свойство равномерной эргодичности, так что распределение случайного вектора $\mathbf{z}_i$ стремится к предельно равномерному распределению с геометрической прогрессией.

Практические реализации могут отличаться количеством генерируемых HR-точек, типом их усреднения (анализа для оценивания центра тяжести), выбором начальной точки для генерирования распределения в допустимой области D_k на k-м шаге алгоритма, выбором точки не случайно равномерно на хорде, а с учетом значений целевой функции в этих точках.

Данный перечень алгоритмов является далеко неполным и дает лишь представление о многообразии подходов для нахождения значения двойственной оценки. Несмотря на доминирующее в настоящее время использование для решения SDP-задач методов внутренней точки, которые характеризуются эффективной как теоретической, так и практической сходимостью, в ряде случаев могут эффективно использоваться другие вычислительные схемы. Наиболее часто приводимые доводы — простота реализации, возможность учета специальной структуры задачи и использования схем разложения в случае квазиблочной структуры, применимость для других типов задач негладкой выпуклой оптимизации, высокая практическая сходимость при небольших размерах задачи и др.

3. КОНИЧЕСКАЯ РЕГУЛЯРИЗАЦИЯ В ЗАДАЧАХ ВЫПУКЛОЙ ОПТИМИЗАЦИИ

Рассмотрим подход, в котором для задачи выпуклого программирования с ограничениями формируется эквивалентная задача безусловной оптимизации, целевая функция которой принимает конечные значения при любых значениях переменных. Целевая функция исходной задачи может быть не определена на границе и вне допустимой области. Предлагаемый подход обобщает результаты, изложенные в [6], и позволяет преодолевать сложности, возникающие при решении задачи (4).

Рассмотрим задачу выпуклого программирования: найти

$$\psi^* = \inf \psi(\mathbf{u}) \tag{7}$$

при ограничениях

$$h(\mathbf{u}) \le 0, \tag{8}$$

где $\mathbf{u} \in R^n$, $\psi, h: R^n \to R \cup \{+\infty\}$ — выпуклые замкнутые функции.

Обозначим $C = \{\mathbf{u} \in R^n : h(\mathbf{u}) \le 0\}$. Если множество C описывается совокупностью ограничений, т.е. $C = \{\mathbf{u} \in R^n : h_i(\mathbf{u}) \le 0,\ i = 1, \ldots, m\}$, полагаем $h(\mathbf{u}) = \max\{h_i(\mathbf{u}) : i = 1, \ldots, m\}$.

В дальнейшем будем считать, что имеют место следующие предположения.

Предположение 1. Выполняются условия: $\mathrm{int}\, C \subseteq \mathrm{dom}\, \psi$; если $\mathbf{u}$ принадлежит границе множества C, то $h(\mathbf{u}) = 0$.

Предположение 2. Задана допустимая точка $\mathbf{u}^0 \in C$ такая, что $h(\mathbf{u}^0) < 0$.

На границе множества C функция ψ может быть не определена (значение функции ψ может быть равно $+\infty$).

Из замкнутости функции ψ следует, что если $\bar{\mathbf{u}}$ принадлежит границе множества C и для любой последовательности $\mathbf{u}^k \in \mathrm{int}\, C$, $k = 1, \ldots$, такой, что $\mathbf{u}^k \to \bar{\mathbf{u}}$ при $k \to +\infty$, выполняется $\lim\limits_{k \to \infty} \psi(\mathbf{u}^k) < +\infty$, тогда $\bar{\mathbf{u}} \in \mathrm{dom}\, \psi$.

В работе [6] предполагалось, что C — ограниченное замкнутое множество, $C \subseteq \mathrm{dom}\, \psi$. Для оценочных задач квадратичной оптимизации множество C может быть неограниченным, на границе множества C значение функции ψ может быть равно $+\infty$. Полученные в [6] результаты непосредственно обобщаются на рассматриваемый случай. Доказательства леммы 1, теорем 1–4 повторяют доказательства соответствующих утверждений в [6].

Пусть задано некоторое число $E < \psi(\mathbf{u}^0)$. Обозначим F надграфик функции ψ на множестве C:

$$F = \{(\lambda, \mathbf{u}) \in R \times C : \lambda \ge \psi(\mathbf{u})\}.$$

Положим $\mathbf{z} = (\lambda, \mathbf{u}), \mathbf{z} \in R \times R^n$. Рассмотрим коническую оболочку $\mathrm{K}(E)$ надграфика F с вершиной в точке $\mathbf{z}_E^0 = (E, \mathbf{u}^0)$:

$$\mathrm{K}(E) = \{\mathbf{v} : \mathbf{v} \in R \times R^n,\ \mathbf{v} = \mathbf{z}_E^0 + \alpha(\mathbf{z} - \mathbf{z}_E^0),\ \alpha \ge 0,\ \mathbf{z} \in F\}.$$

Множество $\mathrm{K}(E)$ выпукло (поскольку выпуклым является множество F), однако оно может быть незамкнутым (например, если множество C не ограничено, а ψ — линейная функция).

Обозначим $\overline{\mathrm{K}}(E)$ замыкание множества $\mathrm{K}(E)$. Множество $\overline{\mathrm{K}}(E)$ может рассматриваться как надграфик некоторой выпуклой функции. Эту функцию обозна-

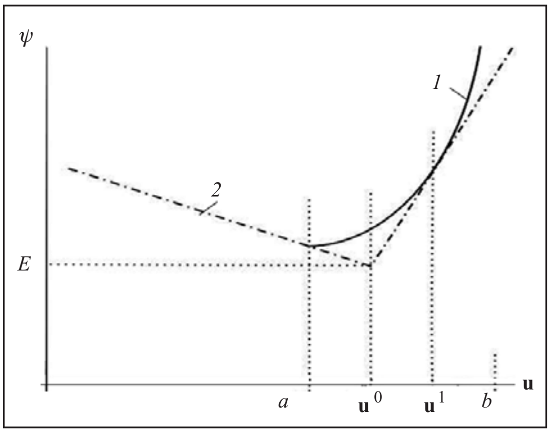

Рис. 1. Коническая аппроксимация функции $\psi(\mathbf{u})$; *1* — функция $\psi(\mathbf{u})$; *2* — функция $\gamma_E(\mathbf{u})$; $a, \mathbf{u}^1$ — точки совпадения значений функций $\psi(\mathbf{u})$ и $\gamma_E(\mathbf{u})$; множество C — отрезок $[a,b]$

чим $\gamma_E(\mathbf{u})$ и назовем конической аппроксимацией функции ψ на множестве C. Функция $\gamma_E(\mathbf{u})$ определена на всем пространстве R^n и принимает конечные значения при любых $\mathbf{u}$. Пример конической аппроксимации функции $\psi(\mathbf{u})$ приведен на рис. 1.

Утверждение 1. Пусть множество C ограничено. Тогда для произвольной точки $\mathbf{u} \in R^n$, $\mathbf{u} \neq \mathbf{u}^0$, на луче, выходящем из точки $\mathbf{u}^0$ и проходящем через $\mathbf{u}$, найдется точка $\bar{\mathbf{u}}$, $\bar{\mathbf{u}} \in C$ (возможно, не одна), такая, что $\psi(\bar{\mathbf{u}}) = \gamma_E(\bar{\mathbf{u}})$. Обозначим $\boldsymbol{\mu}_E(\mathbf{u})$ такую точку, ближайшую к $\mathbf{u}^0$. Если множество C не ограничено, такой точки может не существовать.

Положим

$$\eta_E(\mathbf{u}) = \begin{cases} \|\boldsymbol{\mu}_E(\mathbf{u}) - \mathbf{u}^0\|, & \text{если } \boldsymbol{\mu}_E(\mathbf{u}) \text{ существует,} \\ +\infty & \text{в противном случае;} \end{cases}$$

$$\varphi_E(\mathbf{u}) = \begin{cases} \psi(\mathbf{u}), & \text{если } \|\mathbf{u} - \mathbf{u}^0\| \le \eta_E(\mathbf{u}), \\ \gamma_E(\mathbf{u}), & \text{если } \|\mathbf{u} - \mathbf{u}^0\| > \eta_E(\mathbf{u}). \end{cases}$$

Лемма 1. Пусть точка $\mathbf{u} \in R^n$, $\mathbf{u} \neq \mathbf{u}^0$, такая, что существует точка $\boldsymbol{\mu}_E(\mathbf{u})$, тогда

$$\gamma_E(\mathbf{u}) = E + (\psi(\boldsymbol{\mu}_E(\mathbf{u})) - E) \frac{\|\mathbf{u} - \mathbf{u}^0\|}{\|\boldsymbol{\mu}_E(\mathbf{u}) - \mathbf{u}^0\|}. \tag{9}$$

Рассмотрим задачу: найти

$$\varphi_E^* = \inf \{\varphi_E(\mathbf{u}) : \mathbf{u} \in R^n\}. \tag{10}$$

Теорема 1. Пусть $E \le \psi^*$, тогда $\varphi_E^* = \psi^*$.

Теорема 2. Пусть выполняются предположения 1, 2 и $E < \psi(\mathbf{u}^0)$. Тогда $\varphi_E : R^n \to R$ — выпуклая функция.

Задачу (10) назовем конической регуляризацией исходной задачи (7), (8).

Для использования рассматриваемого подхода в алгоритмах решения оптимизационных задач должны быть определены эффективные процедуры вычисления значений функции $\varphi_E(\mathbf{u})$ и ее субградиентов (ε-субградиентов).

Обозначим $\psi'(\mathbf{u}, \mathbf{p})$ производную функции ψ в точке $\mathbf{u} \in C$ по направлению $\mathbf{p}$. Пусть зафиксирована некоторая точка $\mathbf{u}^1$, положим $\mathbf{p} = \dfrac{\mathbf{u}^1 - \mathbf{u}^0}{\|\mathbf{u}^1 - \mathbf{u}^0\|}$, $t^* = \|\boldsymbol{\mu}_E(\mathbf{u}^1) - \mathbf{u}^0\|$; если точки $\boldsymbol{\mu}_E(\mathbf{u}^1)$ не существует, полагаем $t^* = +\infty$.

Следуя работе [6], можно показать, что

$$t^* = \sup\left\{ t : \frac{\psi(\mathbf{u}^0 + t\,\mathbf{p}) - E}{t} > \psi'(\mathbf{u}^0 + t\,\mathbf{p}, \mathbf{p}),\ t \ge 0,\ \mathbf{u}^0 + t\,\mathbf{p} \in \operatorname{int} C \right\}. \quad (11)$$

Теорема 3. Пусть точка $\mathbf{u}^1$ такая, что $\overline{\mathbf{u}} = \boldsymbol{\mu}_E(\mathbf{u}^1)$ — внутренняя точка множества C. Тогда в точке $\overline{\mathbf{u}}$ существует субградиент $\overline{\mathbf{g}}$ функции ψ, для которого выполняется равенство

$$\psi(\overline{\mathbf{u}}) - E = \langle \overline{\mathbf{g}}, \overline{\mathbf{u}} - \mathbf{u}^0 \rangle, \quad (12)$$

и вектор $\overline{\mathbf{g}}$ есть субградиент функции γ_E в точке $\mathbf{u}^1$ (в точке $\overline{\mathbf{u}}$).

Теорема 4. Пусть выполняются предположения 1, 2, точка $\mathbf{u}^1$ — такая, что $\overline{\mathbf{u}} = \boldsymbol{\mu}_E(\mathbf{u}^1)$ принадлежит границе множества C, тогда

1) $\overline{\mathbf{u}} \in \operatorname{dom} \psi$,

2) $\langle \mathbf{g}_h(\overline{\mathbf{u}}), \mathbf{u}^0 - \overline{\mathbf{u}} \rangle \ne 0$ и вектор

$$\mathbf{g} = \mathbf{g}_\psi(\overline{\mathbf{u}}) + \frac{E - \psi(\overline{\mathbf{u}}) - \langle \mathbf{g}_\psi(\overline{\mathbf{u}}), \mathbf{u}^0 - \overline{\mathbf{u}} \rangle}{\langle \mathbf{g}_h(\overline{\mathbf{u}}), \mathbf{u}^0 - \overline{\mathbf{u}} \rangle}\, \mathbf{g}_h(\overline{\mathbf{u}})$$

есть субградиент функции $\gamma_E(\mathbf{u})$ в точке $\mathbf{u}^1$ (в точке $\overline{\mathbf{u}}$), где $\mathbf{g}_\psi(\overline{\mathbf{u}})$, $\mathbf{g}_h(\overline{\mathbf{u}})$ — субградиенты функций ψ и h в точке $\overline{\mathbf{u}}$.

При использовании предлагаемого подхода значение ψ^* неизвестно. В связи с этим величина E уточняется в процессе решения задачи (10). При получении допустимого решения $\widetilde{\mathbf{u}}$ задачи (7), (8) такого, что $\psi(\widetilde{\mathbf{u}}) < E$, полагаем $E = \psi(\widetilde{\mathbf{u}}) - B$, где $B > 0$. При фиксированном значении параметра B число таких уточнений будет конечно.

Необходимо отметить, что для вычисления значения функции φ_E в произвольной точке $\mathbf{u}$ должна решаться задача одномерного поиска (11), что в случае общей задачи выпуклого программирования существенно увеличивает трудоемкость по сравнению с вычислением значения функции ψ.

4. КОНИЧЕСКАЯ РЕГУЛЯРИЗАЦИЯ ОЦЕНОЧНОЙ КВАДРАТИЧНОЙ ЗАДАЧИ

Основным содержанием данного раздела является анализ свойств оценочной квадратичной задачи (4), что позволяет сформировать эффективный алгоритм одномерного поиска для задачи (11). Трудоемкость такого алгоритма оказывается соизмеримой с трудоемкостью вычисления значения функции ψ.

4.1. Задачу (4) представим в следующем виде: найти

$$\psi^* = \inf\,\{\psi(\mathbf{u}) : \mathbf{u} \in \overline{D}\}, \quad (13)$$

где $\overline{D} = \{\mathbf{u} : \lambda_{\max}(\mathbf{A}(\mathbf{u})) \le 0, \mathbf{u} = (u_1, \ldots, u_N),\ u_i \le 0, i = 1, \ldots, m_1\}$. Функция $\psi(\mathbf{u})$ определяется решением задачи (3).

Пусть заданы точка $\mathbf{u}^0$ такая, что $\lambda_{\max}(\mathbf{A}(\mathbf{u}^0)) < 0$, и произвольная точка $\mathbf{u}^1 \in R^N$. Рассмотрим точки $\mathbf{u}(t) = \mathbf{u}^0 + t(\mathbf{u}^1 - \mathbf{u}^0)$, расположенные на прямой, проходящей через точки $\mathbf{u}^0$ и $\mathbf{u}^1$.

Введем обозначения: $\mathbf{A}^0 = \mathbf{A}(\mathbf{u}^0)$, $\mathbf{A}^1 = \mathbf{A}(\mathbf{u}^1)$, $\mathbf{l}^0 = \mathbf{l}(\mathbf{u}^0)$, $\Delta\mathbf{l} = \mathbf{l}(\mathbf{u}^1) - \mathbf{l}(\mathbf{u}^0)$, $c^0 = c(\mathbf{u}^0)$, $\Delta c = c(\mathbf{u}^1) - c(\mathbf{u}^0)$,

$$\mathbf{A}(\mathbf{u}(t)) = \mathbf{A}^0 + t(\mathbf{A}^1 - \mathbf{A}^0),$$

$$\mathbf{l}(\mathbf{u}(t)) = \mathbf{l}^0 + t(\mathbf{l}(\mathbf{u}^1) - \mathbf{l}(\mathbf{u}^0)) = \mathbf{l}^0 + t\Delta\mathbf{l},$$

$$c(\mathbf{u}(t)) = c^0 + t(c(\mathbf{u}^1) - c(\mathbf{u}^0)) = c^0 + t\Delta c.$$

Для $\mathbf{u}(t) = \mathbf{u}^0 + t(\mathbf{u}^1 - \mathbf{u}^0)$, $\mathbf{u}(t) \in \{\mathbf{u}: \lambda_{\max}(\mathbf{A}(\mathbf{u})) < 0\}$ система уравнений (5) принимает вид

$$(\mathbf{A}^0 + t(\mathbf{A}^1 - \mathbf{A}^0))\mathbf{x} = -\frac{1}{2}(\mathbf{l}^0 + t\Delta\mathbf{l}). \tag{14}$$

Известно (см., например, [29]), что для вещественных симметричных матриц $-\mathbf{A}^0$, $\mathbf{A}^1$ при условии, что матрица $-\mathbf{A}^0$ положительно определена, существует невырожденная матрица $\mathbf{T}$ такая, что

$$\mathbf{T}'(-\mathbf{A}^0)\mathbf{T} = \mathbf{I},\ \mathbf{T}'\mathbf{A}^1\mathbf{T} = \mathbf{B}, \tag{15}$$

где $\mathbf{I}$ — единичная, $\mathbf{B}$ — диагональная матрицы.

Пусть матрица $\mathbf{T}$ удовлетворяет условию (15). Полагая $\mathbf{x} = \mathbf{T}\mathbf{y}$, уравнение (14) приведем к виду $(-\mathbf{I} + t(\mathbf{B}+\mathbf{I}))\mathbf{y} = -\frac{\mathbf{1}}{\mathbf{2}}\mathbf{T}'(\mathbf{l}^0 + t\Delta\mathbf{l})$ или

$$(-1 + (B_{ii}+1)t)\,y_i = -\frac{1}{2}(\eta_i^0 + t\Delta\eta_i),\ i = 1, \ldots, n,$$

где $\boldsymbol{\eta}^0 = \mathbf{T}'\mathbf{l}^0$, $\Delta\boldsymbol{\eta} = \mathbf{T}'\Delta\mathbf{l}$, а η_i^0, $\Delta\eta_i$, y_i, $i = 1, \ldots, n$, — компоненты векторов $\boldsymbol{\eta}^0$, $\Delta\boldsymbol{\eta}$, $\mathbf{y}$.

При выполнении условия $\mathbf{u}(t) \in \{\mathbf{u}: \lambda_{\max}(\mathbf{A}(\mathbf{u})) < 0\}$ имеем простые правила определения вектора $\mathbf{x}(\mathbf{u}(t))$ и вычисления функции $\psi(t) = \psi(\mathbf{u}(t))$:

$$\mathbf{x}(\mathbf{u}(t)) = \mathbf{T}\mathbf{y}(t),\ \text{где}\ y_i(t) = -\frac{(\eta_i^0 + t\Delta\eta_i)}{2(-1 + (B_{ii}+1)t)},\ i = 1, \ldots, n, \tag{16}$$

$$\psi(t) = \psi(\mathbf{u}(t)) = -\frac{1}{4}\sum_{i=1}^{n}\frac{(\eta_i^0 + t\Delta\eta_i)^2}{(-1 + (B_{ii}+1)t)} + c^0 + t\Delta c. \tag{17}$$

Для градиента $\mathbf{g}_\psi(\mathbf{u})$ функции $\psi(\mathbf{u})$ в точке $\mathbf{x} = \mathbf{x}(\mathbf{u}(t))$ в соответствии с (6) имеем

$$\mathbf{g}_\psi(\mathbf{u}(t)) = \mathbf{K}(\mathbf{x}(\mathbf{u}(t))).$$

Положим

$$t_{\max} = \min\{1/(B_{ii}+1): B_{ii} > -1,\ i = 1, \ldots, n\},$$

$$\text{если}\ \{i: B_{ii} > -1,\ i = 1, \ldots, n\} = \varnothing\},\ \text{то}\ t_{\max} = \infty,$$

$$t_{\min} = \max\{1/(B_{ii}+1): B_{ii} < -1,\ i = 1, \ldots, n\},$$

$$\text{если}\ \{i: B_{ii} < -1,\ i = 1, \ldots, n\} = \varnothing,\ \text{то}\ t_{\min} = -\infty.$$

Нетрудно проверить справедливость следующих утверждений.

Утверждение 2. Пусть $t_{\min} < t < t_{\max}$, тогда $\lambda_{\max}(\mathbf{A}(\mathbf{u}(t))) < 0$. Если $t_{\max} < \infty$, то $\lambda_{\max}(\mathbf{A}(\mathbf{u}(t_{\max}))) = 0$.

Аналогично имеет место утверждение для $t_{\min}$.

Утверждение 3. Пусть $i^* \in \arg\min \{1/(B_{ii}+1): B_{ii} > -1,\ i = 1, \dots, n\}$, тогда

$$\lim_{t<t_{\max},\, t\to t_{\max}} \frac{(\eta^0_{i^*} + t\Delta\eta_{i^*})^2}{(-1+(B_{i^*i^*}+1)t)} = \begin{cases} 0, & \text{если } \eta^0_{i^*} + t_{\max}\Delta\eta_{i^*} = 0, \\ -\infty & \text{в противном случае.} \end{cases}$$

Аналогичное утверждение может быть сформулировано для

$$i^* \in \arg\max \{1/(B_{ii}+1):\ B_{ii} < -1,\ i = 1, \dots, n\}.$$

Обозначим $I^* = \arg\min \{1/(B_{ii}+1): B_{ii} > -1,\ i = 1, \dots, n\}$.

Теорема 5. Пусть $t_{\max} < \infty$, $\lim\limits_{t<t_{\max},\, t\to t_{\max}} \psi(t) < \infty$, т.е. $\eta^0_i + t_{\max}\Delta\eta_i = 0,\ i \in I^*$.

Положим

$$y_i(t_{\max}) = -\frac{\eta^0_i + t_{\max}\Delta\eta_i}{2(-1+(B_{ii}+1)t_{\max})},\ i \notin I^*, \tag{18}$$

для $i \in I^*$ значения $y_i(t_{\max})$ определим произвольным образом. Тогда вектор $\mathbf{x}(\mathbf{u}(t_{\max})) = \mathbf{T}\mathbf{y}(t_{\max})$ есть решение задачи (13) при $\mathbf{u} = \mathbf{u}(t_{\max})$.

Доказательство. С учетом введенных обозначений приведем функцию Лагранжа $L(\mathbf{x}, \mathbf{u}(t_{\max}))$ к виду

$$L(\mathbf{x}, \mathbf{u}(t_{\max})) = \langle (\mathbf{A}^0 + t_{\max}(\mathbf{A}^1 - \mathbf{A}^0))\mathbf{x}, \mathbf{x}\rangle + \langle (\mathbf{l}^0 + t_{\max}\Delta\mathbf{l}), \mathbf{x}\rangle + c^0 + t_{\max}\Delta c =$$

$$= \langle (-\mathbf{I} + t_{\max}(\mathbf{B}+\mathbf{I}))\mathbf{y}, \mathbf{y}\rangle + \langle (\boldsymbol{\eta}^0 + t_{\max}\Delta\boldsymbol{\eta}), \mathbf{y}\rangle + c^0 + t_{\max}\Delta c.$$

Учитывая, что $(-1 + t_{\max}(B_{ii}+1)) = 0$, $\eta^0_i + t_{\max}\Delta\eta_i = 0$, $i \in I^*$, т.е. от переменных $y_i, i \in I^*$, функция $L(\mathbf{x}, \mathbf{u}(t_{\max}))$ не зависит, и записывая условия оптимальности относительно остальных переменных, получаем утверждение теоремы. ■

Таким образом, теорема 5 определяет решения задачи (3) в точках на границе ($\lambda_{\max}(\mathbf{A}(\mathbf{u})) = 0$) множества $\overline{D}$, в которых функция $\psi(\mathbf{u})$ принимает конечные значения. Покажем, повторяя доказательство леммы 1.1 из [30], что такие решения, в свою очередь, определяют субградиенты $\mathbf{g}_\psi(\mathbf{u}(t_{\max}))$ функции $\psi(\mathbf{u})$:

$$\mathbf{g}_\psi(\mathbf{u}(t_{\max})) = \mathbf{K}(\mathbf{x}(\mathbf{u}(t_{\max}))). \tag{19}$$

Здесь, как и в (6), $\mathbf{K}(\mathbf{x})$ — вектор, компонентами которого являются функции $K_i(\mathbf{x})$, $i = 1, \dots, N$, формирующие ограничения задачи (1), (2).

Нетрудно увидеть, что для произвольной точки $u \in R^N$ имеет место

$$\psi(\mathbf{u}) - \psi(\mathbf{u}(t_{\max})) \ge L(\mathbf{x}(\mathbf{u}(t_{\max})), \mathbf{u}) - L(\mathbf{x}(\mathbf{u}(t_{\max})), \mathbf{u}(t_{\max})).$$

Действительно, если $\psi(\mathbf{u})$ принимает конечное значение, то

$$\psi(\mathbf{u}) - \psi(\mathbf{u}(t_{\max})) = L(\mathbf{x}(\mathbf{u}), \mathbf{u}) - L(\mathbf{x}(\mathbf{u}(t_{\max})), \mathbf{u}(t_{\max})) \ge$$

$$\ge L(\mathbf{x}(\mathbf{u}(t_{\max})), \mathbf{u}) - L(\mathbf{x}(\mathbf{u}(t_{\max})), \mathbf{u}(t_{\max})).$$

Здесь $\mathbf{x}(\mathbf{u})$ — решение задачи (3) при заданном значении вектора $\mathbf{u}$. Если $\psi(\mathbf{u}) = \infty$, то неравенство очевидно.

Далее,

$$L(\mathbf{x}(\mathbf{u}(t_{\max})), \mathbf{u}) - L(\mathbf{x}(\mathbf{u}(t_{\max})), \mathbf{u}(t_{\max})) = \sum_{i=1}^{N} (u_i - u_i(t_{\max}))K_i(\mathbf{x}(\mathbf{u}(t_{\max}))) =$$

$$= \sum_{i=1}^{N} (u_i - u_i(t_{\max})) K_i(\mathbf{x}(\mathbf{u}(t_{\max}))) = \langle \mathbf{u} - \mathbf{u}(t_{\max}), \mathbf{K}(\mathbf{x}(\mathbf{u}(t_{\max}))) \rangle,$$

откуда следует (19).

4.2. При использовании конической регуляризации для задачи (13) предполагается заданной начальная допустимая точка $\mathbf{u}^0$ такая, что $\lambda_{\max}(\mathbf{A}(\mathbf{u}^0)) < 0$. Для произвольной точки $\mathbf{u}^1 \in R^N$ необходимо найти точку $\boldsymbol{\mu}_E(\mathbf{u}^1)$ на луче, выходящем из точки $\mathbf{u}^0$ и проходящем через точку $\mathbf{u}^1$, в которой выполняется равенство $\psi(\boldsymbol{\mu}_E(\mathbf{u}^1)) = \gamma_E(\boldsymbol{\mu}_E(\mathbf{u}^1))$. Здесь $\gamma_E(\mathbf{u})$ — коническая аппроксимация функции ψ на множестве $\overline{D}$. Для поиска такой точки необходимо решить задачу одномерного поиска (11), которая с учетом (17) может быть представлена в виде

$$t^* = \sup\left\{ t : \frac{\psi(t) - E}{t} > \psi'(t),\ 0 \le t < t_{\max},\ u_i(t) \le 0,\ i = 1, \ldots, m_1 \right\}. \qquad (20)$$

Функция $\psi(t)$ является простой, и задача (20) может быть решена достаточно эффективно. В случае $t^* = \infty$ функция $\gamma_E(\mathbf{u})$ не используется при регуляризации исходной задачи (13) вдоль прямой, проходящей через точки $\mathbf{u}^0$ и $\mathbf{u}^1$.

Наиболее трудоемким при решении задачи (20) является построение матрицы $\mathbf{T}$, удовлетворяющей условию (15). Заметим, что при вычислении значения функции $\psi(\mathbf{u})$ необходимо проверять условие $\lambda_{\max}(\mathbf{A}(\mathbf{u})) < 0$, что соизмеримо по трудоемкости с построением матрицы $\mathbf{T}$. Таким образом, трудоемкости вычисления значений функций $\psi(\mathbf{u})$ и $\varphi_E(\mathbf{u})$ оказываются соизмеримыми.

Рассмотрим правила вычисления субградиентов функции $\gamma_E(\mathbf{u})$.

Если $t^* < t_{\max} < \infty$, $u_i(t^*) < 0$, $i = 1, \ldots, m_1$, то субградиент функции $\gamma_E(\mathbf{u})$ в точке $\mathbf{u}(t^*)$ определяется в соответствии с теоремой 3 решением $\mathbf{x}(\mathbf{u}(t^*))$, которое задается соотношением (16).

Если $t^* < t_{\max} < \infty$, $u_i(t^*) = 0$, $i \in \{1, \ldots, m_1\}$ (точка $\mathbf{u}(t^*)$ лежит на границе множества $\{\mathbf{u} : u_i \le 0, i = 1, \ldots, m_1\}$), то субградиент функции $\gamma_E(\mathbf{u})$ в точке $\mathbf{u}(t^*)$ определяется в соответствии с теоремой 4.

В случае $t^* = t_{\max} < \infty$ субградиент функции $\gamma_E(\mathbf{u})$ в точке $\mathbf{u}(t^*)$ определяется в соответствии со следующей теоремой.

Теорема 6. Пусть $t^* = t_{\max} < \infty$, $u_i(t^*) < 0$, $i = 1, \ldots, m_1$. Тогда существует вектор $\mathbf{y}^* = (y_i^*, i = 1, \ldots, n)$, удовлетворяющий квадратичному уравнению

$$-\langle \mathbf{y}, \mathbf{y} \rangle + \langle \mathbf{T}' \mathbf{l}^{\mathbf{0}}, \mathbf{y} \rangle + c^0 = E \qquad (21)$$

и соотношениям (18) теоремы 5. При этом вектор $\mathbf{x}^* = \mathbf{T}\mathbf{y}^*$ является решением задачи (3) при $\mathbf{u} = \mathbf{u}(t^*)$, а вектор $(K_i(\mathbf{x}^*))_{i=1}^N$ является субградиентом функции $\gamma_E(\mathbf{u})$ в точке $\mathbf{u}(t^*)$. Здесь матрица $\mathbf{T}$ удовлетворяет условию (15).

Доказательство. Из теоремы 4 следует, что если для субградиента $\mathbf{g}_\psi$ функции ψ в точке $\mathbf{u}(t^*)$ выполняется равенство

$$\psi(\mathbf{u}(t^*)) - E = \langle \mathbf{g}_\psi, \mathbf{u}(t^*) - \mathbf{u}^0 \rangle, \qquad (22)$$

то вектор $\mathbf{g}_\psi$ есть субградиент функции γ_E в этой точке. Субградиенты $\mathbf{g}_\psi$ в точке $\mathbf{u}(t^*)$ определяются решениями $\mathbf{x}^*$ задачи (3) при $\mathbf{u}=\mathbf{u}(t^*)$, т.е. $\mathbf{g}_\psi=\mathbf{K}(\mathbf{x}^*)=(K_i(\mathbf{x}^*))_{i=1}^N$, откуда

$$\psi(\mathbf{u}(t^*))-E=\langle \mathbf{K}(\mathbf{x}^*),\mathbf{u}(t^*)-\mathbf{u}^0\rangle. \tag{23}$$

Решения $\mathbf{x}^*$ в соответствии с теоремой 5 могут быть представлены в виде $\mathbf{x}^*=\mathbf{T}\mathbf{y}^*$, где матрица $\mathbf{T}$ и вектор $\mathbf{y}^*$ удовлетворяют условиям теоремы 5.

Пусть $\mathbf{x}^*$ — решение задачи (3) при $\mathbf{u}=\mathbf{u}(t^*)$, для которого выполняется условие (23). Нетрудно видеть, что $\langle \mathbf{g}_\psi,\mathbf{u}(t^*)-\mathbf{u}^0\rangle=\langle \mathbf{K}(\mathbf{x}^*),\mathbf{u}(t^*)-\mathbf{u}^0\rangle=$ $=L(\mathbf{x}^*,\mathbf{u}(t^*))-L(\mathbf{x}^*,\mathbf{u}^0)=\psi(\mathbf{u}(t^*))-L(\mathbf{x}^*,\mathbf{u}^0)$. С учетом (22) получаем $L(\mathbf{x}^*,\mathbf{u}^0)=E$. Кроме того, имеет место

$$L(x^*,u^0)=\langle \mathbf{A}(\mathbf{u}^0)\mathbf{x}^*,\mathbf{x}^*\rangle+\langle \mathbf{l}(\mathbf{u}^0),\mathbf{x}^*\rangle+c(\mathbf{u}^0)=\langle \mathbf{A}^0\mathbf{x}^*,\mathbf{x}^*\rangle+\langle \mathbf{l}^0,\mathbf{x}^*\rangle+c^0.$$

Проведя замену $\mathbf{x}^*=\mathbf{T}\mathbf{y}^*$ и учитывая, что матрица $\mathbf{T}$ удовлетворяет условию (15), получаем $L(\mathbf{x}^*,\mathbf{u}^0)=-\langle \mathbf{y}^*,\mathbf{y}^*\rangle+\langle \mathbf{T}'\mathbf{l}^{\mathbf{0}},\mathbf{y}^*\rangle+c^0$. Отсюда следует выполнение условия (21).

Уравнение (21) при дополнительных соотношениях (18) теоремы 5 разрешимо не при любых значениях параметра E. В частности, если $E>\psi(\mathbf{u}^0)$, то уравнение (21) решений не имеет. По построению конической аппроксимации $\gamma_E(\mathbf{x})$ величина E выбирается так, что $E<\psi(\mathbf{u}^0)$. Можно показать, что если уравнение (21) решений не имеет, то найдется точка $t'<t_{\max}$, в которой будет выполняться условие (12). Последнее противоречит предположению о том, что $t^*=t_{\max}$. Теорема доказана. ■

Полученные результаты позволяют формировать эффективные алгоритмы для реализации конической регуляризации оценочной квадратичной задачи.

ЗАКЛЮЧЕНИЕ

В настоящей статье для вычисления оценки оптимального значения задачи квадратичной оптимизации предложено использовать коническую регуляризацию оценочной задачи. Такой подход позволяет построить эквивалентную задачу безусловной оптимизации, целевая функция которой определена на всем пространстве переменных задачи и удовлетворяет условию Липшица. Особенностью рассмотренного класса задач является возможность построения эффективных алгоритмов поиска по направлению. Такие алгоритмы являются существенными для использования конической регуляризации. Трудоемкость вычисления функций и субградиентов регуляризованой задачи оказывается соизмеримой с трудоемкостью вычисления функций исходной оценочной задачи.

Полученные результаты полезны при разработке эффективных методов решения оценочных задач квадратичной оптимизации.

СПИСОК ЛИТЕРАТУРЫ

1. Lemarechal C. Lagrangian relaxation. *Computational combinatorial optimization*. 2001. P. 112–156.
2. Шор Н.З., Стеценко С.И. Квадратичные экстремальные задачи и недифференцируемая оптимизация. Киев: Наукова думка, 1989. 204 с.

3. Boyd S., Vandenberghe L. Semidefinite programming relaxations of non-convex problems in control and combinatorial optimization. *Communications, Computation, Control, and Signal Processing*. New York: Springer, 1997. C. 279–287.

4. Anstreicher K., Wolkowicz H. On Lagrangian relaxation of quadratic matrix constraints. *SIAM Journal on Matrix Analysis and Applications*. 2000. Vol. 22, N 1. C. 41–55.

5. Berezovskyi O.A. Exactness criteria for SDP-relaxations of quadratic extremum problems. *Cybernetics and Systems Analysis*. 2016. Vol. 52, N 6. P. 915–920.

6. Лаптин Ю.П., Бардадым Т.А. Некоторые подходы к регуляризации нелинейных задач оптимизации. *Проблемы управления и информатики*. 2011. № 3. С. 57–68.

7. Лаптин Ю.П. Коническая регуляризация в задачах квадратичной оптимизации. *Компьютерная математика*. 2016. № 2. С. 129–141.

8. Поляк Б.Т., Щербаков П.С. Рандомизированный метод решения задач полуопределенного программирования. *Стохастическая оптимизация в информатике*. 2006. Т. 2, № 1. С. 38–70.

9. Geoffrion A.M. Lagrangian relaxation for integer programming. *Math. Programming Study*. 1974. Vol. 2. P. 82–114.

10. Grötschel M., Lovász L, Schrijver A. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*. 1981. Vol. 1, N 2. P. 169–197.

11. Шор Н.З., Давыдов А.С. О методе получения оценок в квадратичных экстремальных задачах с булевыми переменными. *Кибернетика*. 1985. № 2. С. 48–50.

12. Shor N.Z., Stetsyuk P.I. Modified *r*-algorithm to find the global minimum of polynomial functions. *Cybernetics and Systems Analysis*. 1997. Vol. 33, N 4. P. 482–497.

13. Shor N.Z., Berezovskyi O.A. Using the method of dual quadratic solutions to solve systems of polynomial equations in the complex domain. *Cybernetics and Systems Analysis*. 1994. Vol. 30, N 5. P. 686–692.

14. Berezovskyi O.A., Stetsyuk P.I. An approach to determining Shor's dual quadratic estimates. *Cybernetics and Systems Analysis*. 2008. Vol. 44, N 2. P. 225–233.

15. Шор Н.З., Журбенко Н.Г. Метод минимизации, использующий операцию растяжения пространства в направлении разности двух последовательных градиентов. *Кибернетика*. 1971. № 3. С. 51–59.

16. Vandenberghe L., Boyd S. Semidefinite programming. *SIAM Reviev*. 1996. Vol. 38, N 1. P. 49–95.

17. Nesterov Yu., Nemirovskii A. Interior-points polynomial algorithm in convex programming. Philadelphia: SIAM, 1994. 403 p.

18. Зоркальцев В.И., Пержабинский С.М. Алгоритмы внутренних точек в линейном и нелинейном программировании. *Омский научный вестник*. 2013. Vol. 1, N 117. C. 25–28.

19. Жадан В.Г., Орлов А.А. Допустимый двойственный метод внутренней точки для линейной задачи полуопределенного программирования. *Автоматика и телемеханика*. 2012. № 2. С. 25–40.

20. Евтушенко Ю.Г., Жадан В.Г. Двойственные барьерно-проективные и барьерно-ньютоновские методы для линейного программирования. *Журн. вычисл. математики и мат. физики*. 1994. Т. 36, № 7. С. 30–45.

21. Бабынин М.С., Жадан В.Г. Прямой метод внутренней точки для линейной задачи полуопределенного программирования. *Журн. вычисл. математики и мат. физики*. 2008. Т. 48, № 10. С. 1780–1801.

22. Muramatsu M. Affine scaling algorithm fails for semidefinite programming. *Mathematical programming*. 1998. Vol. 83. P. 393–406.

23. Lasserre J.B. Linear programming with positive semi-definite matreces. *Math. Problems in Engineering*. 1996. Vol. 2, N 6. P. 499–522.

24. Pataki G. Cone-LP's and semidefinite programs: Geometry and simplex-type method. *Proc. Conf. on Integer Programming and Combinatorial Optimization (IPCO-5)*. Vancouver, 1996. P. 1–13.

25. Жадан В.Г. Об одном варианте допустимого аффинно-масштабирующего метода для полуопределенного программирования. *Тр. Ин-та математики и механики УрО РАН*. 2014. Т. 20, № 2. С. 145–160.

26. Жадан В.Г. Об одном варианте симплекс-метода для линейной задачи полуопределенного программирования. *Тр. Ин-та математики и механики УрО РАН*. 2015. Т. 21, № 3. С. 117–127.

27. Smith R.L. Efficient Monte Carlo procedures for generating points uniformly distributed over bounded regions. *Operations Research*. 1984. Vol. 32. P. 1296–1308.

28. Lovász L. Hit-and-run mixes fast. *Math. Progr*. 1999. Vol. 86. P. 443–461.

29. Беллман Р. Введение в теорию матриц. Москва: Наука, 1976. 352 с.

30. Михалевич В.С., Трубин В.А., Шор Н.З. Оптимизационные задачи производственно-транспортного планирования. Москва: Наука, 1986. 260 с.

Надійшла до редакції 03.05.2017

Ю.П. Лаптін, О.А. Березовський

ВИКОРИСТАННЯ КОНІЧНОЇ РЕГУЛЯРИЗАЦІЇ ПРИ ОБЧИСЛЕННІ ЛАГРАНЖЕВИХ ОЦІНОК У ЗАДАЧАХ КВАДРАТИЧНОЇ ОПТИМІЗАЦІЇ

Анотація. Для неопуклих задач квадратичної оптимізації розглянуто обчислення оцінок значень глобальних екстремумів на основі лагранжевої релаксації початкових задач. На границі допустимої області оціночної задачі функції задачі є розривними, погано обумовленими, що накладає певні вимоги на обчислювальні алгоритми. Для урахування зазначених особливостей розроблено новий підхід, який базується на використанні конічних регуляризацій опуклих задач оптимізації. Він дозволяє побудувати еквівалентну задачу безумовної оптимізації, цільова функція якої визначена на всьому просторі змінних задачі і задовольняє умові Ліпшиця.

Ключові слова: задача квадратичної оптимізації, лагранжева релаксація, умова невід’ємної визначеності матриці, конічна регуляризація.

Yu.P. Laptin, O.A. Berezovskyi

USING CONICAL REGULARIZATION IN CALCULATING LAGRANGIAN ESTIMATES IN QUADRATIC OPTIMIZATION PROBLEMS

Abstract. For nonconvex quadratic optimization problems, calculation of global extreme value estimates on the basis of Lagrangian relaxation of the original problems is considered. On the boundary of the feasible region of the estimation problem, the functions of the problem are discontinuous, ill-conditioned, which imposes certain requirements on the computational algorithms. The paper presents a new approach taking into account these features, based on the use of conical regularizations of convex optimization problems. It makes it possible to construct an equivalent unconditional optimization problem, whose objective function is defined on the entire space of problem variables and satisfies the Lipschitz condition.

Keywords: quadratic optimization problem, Lagrangian relaxation, condition of non-negative definiteness of the matrix, conical regularization.

Лаптин Юрий Петрович,
доктор физ.-мат. наук, старший научный сотрудник Института кибернетики им. В.М. Глушкова НАН Украины, Киев, e-mail: yu.p.laptin@gmail.com.

Березовский Олег Анатольевич,
кандидат физ.-мат. наук, старший научный сотрудник Института кибернетики им. В.М. Глушкова НАН Украины, Киев, e-mail: o.a.berezovskyi@gmail.com.