КОМПЬЮТЕРНАЯ МАТЕМАТИКА

Теория и методы оптимизации

УДК 519.8

Ю.П. ЛАПТИН

КОНИЧЕСКАЯ РЕГУЛЯРИЗАЦИЯ В ЗАДАЧАХ КВАДРАТИЧНОЙ ОПТИМИЗАЦИИ

Рассматриваются вопросы вычисления оценок оптимальных значений невыпуклых задач квадратичной оптимизации на основе лагранжевых релаксаций исходной задачи. На границе допустимой области оценочной задачи функции задачи могут быть разрывны, плохо обусловлены, что усложняет разработку вычислительных алгоритмов. В работе приводятся новые подходы преодоления указанных проблем, основанные на использовании конических регуляризаций выпуклых задач оптимизации.

Введение. К невыпуклым задачам квадратичной оптимизации (максимизации квадратичной формы при квадратичных ограничениях) могут быть сведены самые разные многоэкстремальные задачи. Для вычисления оценок оптимальных значений таких задач применяются лагранжевы релаксации (см., например, [1, 2]) и SDP-релаксации (релаксации к задачам полуопределенного программирования) [3], которые в определенном смысле эквивалентны. При использовании лагранжевых релаксаций формируются оценочные параметрические матричные задачи в пространстве двойственных переменных. Функции оценочных задач непрерывно дифференцируемы и принимают конечные значения в областях отрицательной определенности матриц функций Лагранжа исходных квадратичных задач. На границе области отрицательной определенности эти функции разрывны (в области положительной определенности принимают значения $+\infty$), а вблизи границы матрицы плохо обусловлены (причем количество близких к нулю собственных чисел может приближаться к размерности пространства), что приводят к необходимости учета этих особенностей при разработке вычислительных алгоритмов решения оценочных задач.

В работе приводятся новые подходы преодоления указанных проблем, основанные на использовании конических регуляризаций выпуклых задач оптимизации.

В пункте 1 в соответствии с [1, 2] приводятся постановки квадратичных оптимизационных задач и формулируются оценочные задачи. В пункте 2 формулируются новые результаты, касающиеся поиска по направлению в оценочных задачах. Эти результаты идейно близки алгоритму «граничного оракула», введенному в [4]. В пунктах 3, 4 приводится обобщение понятий конической регуляризации для рассматриваемого класса задач, изложенных ранее в [5]. Рассматриваются эффективные алгоритмы вычисления функций и субградиентов регуляризованой задачи.

1. Квадратичные оптимизационные задачи и оценки их оптимальных значений

Постановка задачи

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

при ограничениях

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

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, $c_i \in R$, $i \in \{0, 1, ..., N\}$.

Задача (1) – (2) – многоэкстремальная в общем случае, для вычисления оценки оптимального значения в [1] предложено использовать лагранжеву релаксацию этой задачи. Обозначим $U^- = \{\mathbf{u} = (u_1,...,u_N) : u_i \le 0,\ i = 1,...,m_1\}$.

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

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

$$\mathbf{A}(\mathbf{u}) = \mathbf{A}_0 + \sum_{i=1}^{N} u_i \mathbf{A}_i\ ,\ \mathbf{l}(\mathbf{u}) = \mathbf{l}_0 + \sum_{i=1}^{N} u_i \mathbf{l}_i\ ,\ c(\mathbf{u}) = c_0 + \sum_{i=1}^{N} u_i c_i\ . \tag{4}$$

Пусть 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 $(\bar{D})$ – подмножества R^N, состоящие из таких $\mathbf{u} \in U^-$, что $\mathbf{A}(\mathbf{u})$ является отрицательно определенной (соответственно – неположительно определенной) матрицей. Если $u \notin \bar{D}$, то $\psi(\mathbf{u}) = +\infty$.

Справедливо соотношение [2]

$$\psi^* = \inf_{\mathbf{u} \in U} \psi(\mathbf{u}) = \inf_{\mathbf{u} \in \bar{D}} \psi(\mathbf{u}). \tag{5}$$

Пусть $\mathbf{u} \in D$, $\mathbf{x}(\mathbf{u})$ – решение задачи (3). Вектор $\mathbf{x}(\mathbf{u})$ – решение системы уравнений [1]

$$2\mathbf{A}(\mathbf{u})\mathbf{x} + \mathbf{l}(\mathbf{u}) = 0\ . \tag{6}$$

Откуда, $\mathbf{x}(\mathbf{u}) = -\frac{1}{2}\mathbf{A}(\mathbf{u})^{-1}\mathbf{l}(\mathbf{u})$,

$$\psi(\mathbf{u}) = -\frac{1}{4}\left\langle \mathbf{A}(\mathbf{u})^{-1}\mathbf{l}(\mathbf{u}), \mathbf{l}(\mathbf{u})\right\rangle + c(\mathbf{u}) = \frac{1}{2}\left\langle \mathbf{x}(\mathbf{u}), \mathbf{l}(\mathbf{u})\right\rangle + c(\mathbf{u}). \quad (7)$$

Обозначим $\lambda_{\max}(\mathbf{A}(u))$ максимальное собственное число матрицы $\mathbf{A}(u)$. Множество $\bar{D}$ может быть представлено в виде $\bar{D} = \left\{\mathbf{u} \in R^N : \lambda_{\max}(\mathbf{A}(\mathbf{u})) \le 0,\ u_i \le 0,\ i = 1,...,m_1\right\}$. Во внутренних точках ($\lambda_{\max}(\mathbf{A}(\mathbf{u})) < 0$) множества $\bar{D}$ функция $\psi(\mathbf{u})$ непрерывно дифференцируема и принимает конечные значения, на границе ($\lambda_{\max}(\mathbf{A}(\mathbf{u})) = 0$) множества $\bar{D}$ функция $\psi(\mathbf{u})$ может принимать как конечные значения, так и значения $+\infty$. Особенности поведения функции $\psi(\mathbf{u})$ вблизи границы множества $\bar{D}$ приводит к существенным проблемам при практическом решении задачи (5).

2. Процедуры одномерного поиска для функции $\psi(\mathbf{u})$

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

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

$$c^0 = c(\mathbf{u}^0),\ \Delta c = c(\mathbf{u}^1) - c(\mathbf{u}^0),$$

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

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

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

Для $\mathbf{u}(t) = \mathbf{u}^0 + t\left(\mathbf{u}^1 - \mathbf{u}^0\right)$, $\mathbf{u}(t) \in D$ соотношение (6) принимает вид

$$\left(\mathbf{A}^0 + t(\mathbf{A}^1 - \mathbf{A}^0)\right)x = -\frac{1}{2}\left(\mathbf{l}^0 + t\Delta\mathbf{l}\right). \quad (8)$$

Известно (см., например, [6]), 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

$$\mathbf{T}'\left(-\mathbf{A}^0\right)\mathbf{T} = \mathbf{I},\ \mathbf{T}'\mathbf{A}^1\mathbf{T} = \mathbf{B}, \quad (9)$$

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

Пусть матрица $\mathbf{T}$ удовлетворяет условию (9). Полагая $\mathbf{x} = \mathbf{T}\mathbf{y}$, уравнение (8) приведем к виду $\left(-\mathbf{I} + t(\mathbf{B} + \mathbf{I})\right)\mathbf{y} = -\frac{1}{2}\mathbf{T}'\left(\mathbf{l}^0 + t\Delta\mathbf{l}\right)$, или

$$\left(-1 + \left(B_{ii} + 1\right)t\right) y_i = -\frac{1}{2}\left(\eta_i^0 + t\Delta\eta_i\right),\ i = 1,\dots,n, \qquad (10)$$

где $\boldsymbol{\eta}^0 = \mathbf{T}'\mathbf{l}^0$, $\Delta\boldsymbol{\eta} = \mathbf{T}'\Delta\mathbf{l}$, η_i^0, $\Delta\eta_i$, $i = 1,\dots,n$ – компоненты векторов $\boldsymbol{\eta}^0$, $\Delta\boldsymbol{\eta}$.

Положим

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

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

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

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

Утверждение 1. Пусть $t_{\min} < t < t_{\max}$, тогда $\mathbf{u}(t) \in D$, $\mathbf{x}(\mathbf{u}(t)) = \mathbf{T}\mathbf{y}(t)$, где $y_i(t) = -\dfrac{\left(\eta_i^0 + t\Delta\eta_i\right)}{2\left(-1 + \left(B_{ii} + 1\right)t\right)}$, $i = 1,\dots,n$.

Если $t_{\max} < \infty$, то $\lambda_{\max}(\mathbf{A}(\mathbf{u}(t_{\max}))) = 0$. Аналогично для $t_{\min}$.

Положим $\psi(t) = \psi(\mathbf{u}(t))$. Для $t_{\min} < t < t_{\max}$ функция $\psi(t)$ может быть представлена в виде

$$\psi(t) = \psi(\mathbf{u}(t)) = -\frac{1}{4}\sum_{i=1}^{n}\frac{\left(\eta_i^0 + t\Delta\eta_i\right)^2}{\left(-1 + \left(B_{ii} + 1\right)t\right)} + c^0 + t\Delta c. \qquad (11)$$

Утверждение 2. Пусть $i^* \in \arg\min\left\{1/\left(B_{ii} + 1\right) :\ B_{ii} > -1,\ i = 1,\dots,n\right\}$, тогда

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

Аналогичное утверждение может быть сформулировано для $i^* \in \arg\max\left\{1/\left(B_{ii} + 1\right) : B_{ii} < -1, i = 1,\dots,n\right\}$.

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

Теорема 1. Пусть $t_{\max} < \infty$, $\lim\limits_{t < t_{\max},\, t \to t_{\max}} \psi(t) < \infty$, т. е. $\eta_i^0 + t_{\max}\Delta\eta_i = 0$, $i \in I^*$.

Положим

$$y_i(t_{\max}) = -\frac{\eta_i^0 + t_{\max}\Delta\eta_i}{2\left(-1+\left(B_{ii}+1\right)t_{\max}\right)},\ i \notin I^*, \tag{12}$$

для $i \in I^*$ значения $y_i(t_{\max})$ определим произвольным образом.

Тогда вектор $\mathbf{x}(\mathbf{u}(t_{\max})) = \mathbf{T}\mathbf{y}(t_{\max})$ – решение задачи (3) при $\mathbf{u} = \mathbf{u}(t_{\max})$.

Доказательство. Учитывая введенные обозначения, приведем функцию Лагранжа $L(\mathbf{x}, \mathbf{u}(t_{\max}))$ к виду

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

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

Учитывая, что $\left(-1+t_{\max}\left(B_{ii}+1\right)\right)=0,\quad \eta_i^0 + t_{\max}\Delta\eta_i = 0,\quad i \in I^*$, 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 1 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 [7], что это в свою очередь определяет субградиенты $\mathbf{g}_\psi(\mathbf{u}(t_{\max}))$ функция $\psi(\mathbf{u})$ в таких точках

$$\mathbf{g}_\psi(\mathbf{u}(t_{\max})) = \mathbf{K}(\mathbf{x}(\mathbf{u}(t_{\max})))\,. \tag{13}$$

Здесь $\mathbf{K}(\mathbf{x}(\mathbf{u}(t_{\max})))$ – вектор, компонентами которого являются величины $K_i(\mathbf{x}(\mathbf{u}(t_{\max}))),\ i = 1, \ldots, N$.

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

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

В самом деле, если $\psi(\mathbf{u})$ принимает конечное значение, то $\psi(\mathbf{u}) - \psi(\mathbf{u}(t_{\max})) = L(\mathbf{x}(\mathbf{u}), \mathbf{u}) - L(\mathbf{x}(\mathbf{u}(t_{\max})), \mathbf{u}(t_{\max})) \geq L(\mathbf{x}(\mathbf{u}(t_{\max})), \mathbf{u}) - L(\mathbf{x}(\mathbf{u}(t_{\max})), \mathbf{u}(t_{\max}))$. Здесь $\mathbf{x}(\mathbf{u})$ – решение задачи (3) при заданном значении вектора $\mathbf{u}$. Если $\psi(\mathbf{u}) = \infty$, то неравенство очевидно.

Далее

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

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

откуда следует (13).

Рассмотрим задачу поиска минимума функции $\psi(t)$ (минимума по направлению функции $\psi(\mathbf{u})$), которая имеет вид

$$\inf \left\{ \psi(t) : t_{\min} < t < t_{\max},\ u_i(t) \le 0,\ i = 1,\dots,m_1 \right\}. \tag{14}$$

Задача (14) – простая и может быть эффективно решена. Наиболее трудоемким является построение матрицы $\mathbf{T}$, удовлетворяющей условию (9).

Использование на каждом шаге оптимизационного алгоритма решений задачи (14) для определения величины шага обеспечивает выполнение всех ограничений задачи (5). В качестве точки u^0 должна использоваться точка, сгенерированная на предыдущем шаге. Принципиальные проблемы, возникающие при таком использовании, связаны с тем, что матрица A^0 может быть плохо обусловленной вблизи границы множества $\bar{D}$ или вырожденной на границе множества $\bar{D}$.

3. Коническая регуляризация задач оптимизации

В данном пункте рассматривается 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 [5], и позволяет преодолевать проблемы, возникающие при решении задачи (5).

Рассмотрим задачу выпуклого программирования: найти

$$\psi^* = \inf \psi(\mathbf{u}) \tag{15}$$

при ограничениях

$$h(\mathbf{u}) \le 0, \tag{16}$$

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

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

Предположение 1. $\operatorname{int} C \subseteq \operatorname{dom} \psi$, 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 $\mathbf{u}^k \in \operatorname{int} C$, $k = 1, \ldots$, 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 $\bar{\mathbf{u}} \in \operatorname{dom} \psi$.

В работе [5] предполагалось, что C – компактное множество, $C \subseteq \operatorname{dom} \psi$. Для оценочных задач квадратичной оптимизации множество C может быть неограниченным, на границе множества C значение функции ψ может быть равно $+\infty$.

Полученные в [5] результаты непосредственно обобщаются на рассматриваемый случай.

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

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

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 $\mathbf{z}_E^0 = (E, \mathbf{u}^0)$:

$$\mathrm{K}(E) = \left\{\mathbf{v} : \mathbf{v} \in R \times R^n,\ \mathbf{v} = \mathbf{z}_E^0 + \alpha(\mathbf{z} - \mathbf{z}_E^0), \alpha \geq 0, \mathbf{z} \in F\right\}. \tag{17}$$

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

Обозначим $\bar{\mathrm{K}}(E)$ замыкание множества $\mathrm{K}(E)$. Множество $\bar{\mathrm{K}}(E)$ может рассматриваться как надграфик некоторой выпуклой функции. Эту функцию обозначим $\gamma_E(\mathbf{u})$ и будем называть 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 показан на рисунке.

Утверждение 3. 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 $\psi(\bar{\mathbf{u}}) = \gamma_E(\bar{\mathbf{u}})$. Если множество C не ограничено, такая точка может не существовать.

Обозначим $\boldsymbol{\mu}_E(\mathbf{u})$ такую точку, ближайшую к $\mathbf{u}^0$. Положим

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

$$\varphi_E(\mathbf{u}) = \begin{cases} \psi(\mathbf{u}), \text{ если } \left\|\mathbf{u}-\mathbf{u}^0\right\| \le \eta_E(\mathbf{u}), \\ \gamma_E(\mathbf{u}), \text{ если } \left\|\mathbf{u}-\mathbf{u}^0\right\| > \eta_E(\mathbf{u}). \end{cases} \quad (18)$$

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

$$\gamma_E(\mathbf{u}) = E + (\psi(\boldsymbol{\mu}_E(\mathbf{u})) - E)\frac{\left\|\mathbf{u}-\mathbf{u}^0\right\|}{\left\|\boldsymbol{\mu}_E(\mathbf{u})-\mathbf{u}^0\right\|}. \quad (19)$$

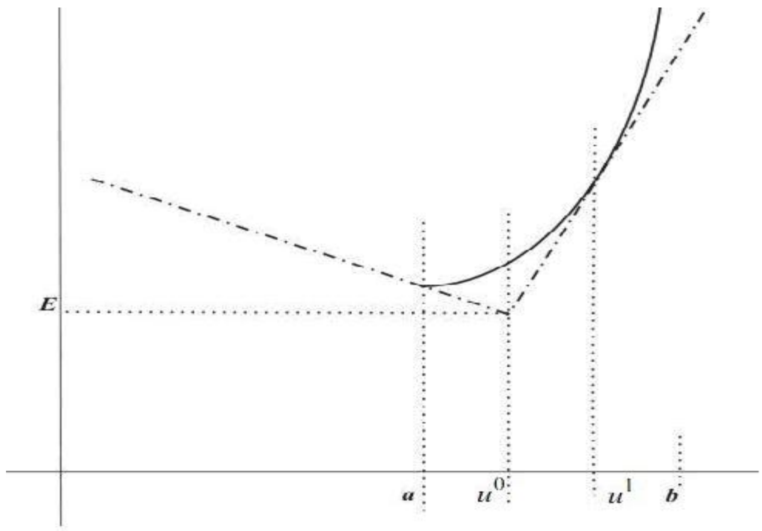

РИСУНОК. Сплошной линией показана функция $\psi(\mathbf{u})$, штрих-пунктирной – $\gamma_E(\mathbf{u})$, в точках $\mathbf{a}$, $\mathbf{u}^1$ значения этих функций совпадают, множество C – отрезок $[\mathbf{a}, \mathbf{b}]$

Рассмотрим задачу: найти

$$\varphi_E^* = \inf\left\{\varphi_E(\mathbf{u}) : \mathbf{u} \in R^n\right\}. \quad (20)$$

Теорема 2. Пусть $E \le \psi^*$, тогда $\varphi_E^* = \psi^*$.

Теорема 3. Пусть выполняются предположения 1, 2 и $E < \psi(x^0)$. Тогда $\varphi_E : R^n \to R$ – выпуклая функция.

Задачу (20) будем называть конической регуляризацией исходной задачи (15) – (16).

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 $\mathbf{p}=\dfrac{\mathbf{u}^1-\mathbf{u}^0}{\left\|\mathbf{u}^1-\mathbf{u}^0\right\|}$, $t^*=\left\|\boldsymbol{\mu}_E(\mathbf{u}^1)-\mathbf{u}^0\right\|$, если точка $\boldsymbol{\mu}_E(\mathbf{u}^1)$ не существует, полагаем $t^*=+\infty$.

Лемма 2. Пусть $t>0$, $\mathbf{u}^0+\mathbf{p}t\in\operatorname{int}C$. Тогда

$$\frac{\psi(\mathbf{u}^0+t\mathbf{p})-E}{t}>\psi'(\mathbf{u}^0+t\mathbf{p},\mathbf{p}),\ \text{если}\ t<t^*, \tag{21}$$

$$\frac{\psi(\mathbf{u}^0+t\mathbf{p})-E}{t}\le-\psi'(\mathbf{u}^0+t\mathbf{p},-\mathbf{p}),\ \text{если}\ t>t^*. \tag{22}$$

Следствие.

$$t^*=\sup\left\{t:\ \frac{\psi(\mathbf{u}^0+t\mathbf{p})-E}{t}>-\psi'(\mathbf{u}^0+t\mathbf{p},-\mathbf{p}),\ \ t\ge 0,\mathbf{u}^0+t\mathbf{p}\in\operatorname{int}C\right\}. \tag{23}$$

Теорема 4. Пусть точка $\mathbf{u}^1$, такая, что $\bar{\mathbf{u}}=\boldsymbol{\mu}_E(\mathbf{u}^1)$ – внутренняя точка *множества* C. Тогда в точке $\bar{\mathbf{u}}$ существует субградиент $\bar{\mathbf{g}}$ функции ψ, для которого выполняется

$$\psi(\bar{\mathbf{u}})-E=\left\langle\bar{\mathbf{g}},\bar{\mathbf{u}}-\mathbf{u}^0\right\rangle \tag{24}$$

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

Теорема 5. Пусть выполняются предположения 1, 2, точка $\mathbf{u}^1$, такая, что $\bar{\mathbf{u}}=\boldsymbol{\mu}_E(\mathbf{u}^1)$ принадлежит границе множества C, тогда

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

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

$$\mathbf{g}=\mathbf{g}_\psi(\bar{\mathbf{u}})+\frac{E-\psi(\bar{\mathbf{u}})-\left\langle\mathbf{g}_\psi(\bar{\mathbf{u}}),\mathbf{u}^0-\bar{\mathbf{u}}\right\rangle}{\left\langle\mathbf{g}_h(\bar{\mathbf{u}}),\mathbf{u}^0-\bar{\mathbf{u}}\right\rangle}\mathbf{g}_h(\bar{\mathbf{u}}) \tag{25}$$

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

При использовании предлагаемого подхода значение ψ^* обычно неизвестно. Поэтому величина E уточняется по ходу решения задачи (20). При получении допустимого решения $\tilde{\mathbf{u}}$ задачи (15) – (16) такого, что $\psi(\tilde{\mathbf{u}}) < E$, полагаем $E = \psi(\tilde{\mathbf{u}}) - B$, 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 (23), что в случае общей задачи выпуклого программирования существенно увеличивает трудоемкость по сравнению с вычислением значения функции ψ.

4. Коническая регуляризация оценочной квадратичной задачи

Подходы, приведенные в пункте 3, будем использовать для решения задачи (5), которую представим в виде

$$\psi^* = \inf\left\{\psi(\mathbf{u}) : \lambda_{\max}(\mathbf{A}(\mathbf{u})) \le 0,\ u_i \le 0,\ i = 1,...,m_1\right\}. \qquad (26)$$

Функция $\psi(\mathbf{u})$ вычисляется в соответствии с соотношением (7), матрица $\mathbf{A}(\mathbf{u})$, вектор $l(\mathbf{u})$ и величина $c(\mathbf{u})$ – в соответствии с (4).

Будем считать заданной точку $\mathbf{u}^0 \in \operatorname{int} D$, такую, что $\lambda_{\max}(\mathbf{A}(\mathbf{u}^0)) < 0$, и число $E < \psi(\mathbf{u}^0)$. При использования конической регуляризации для задачи (26) необходимо для произвольной точки $\mathbf{u}^1 \in R^N$ находить точку $\boldsymbol{\mu}_E(\mathbf{u}^1)$ на луче, выходящем из точки $\mathbf{u}^0$ и проходящем через $\mathbf{u}^1$, 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 $\bar{D}$. Для поиска такой точки должна решаться задача одномерного поиска (23), которая с учетом обозначений, введенных в пункте 2, может быть представлена в виде:

$$t^* = \sup\left\{t : \frac{\psi(t) - E}{t} > \psi'(t),\ 0 \le t < t_{\max},\ u_i(t) \le 0,\ i = 1,...,m_1\right\}. \qquad (27)$$

Функция $\psi(t)$ – простая, задача (27) может быть решена достаточно эффективно. В случае $t^* = \infty$ функция $\gamma_E(\mathbf{u})$ не используется при регуляризации исходной задачи (26) вдоль прямой, проходящей через точки $\mathbf{u}^0$ и $\mathbf{u}^1$.

Как и для задачи (14) наиболее трудоемким при решении задачи (27) является построение матрицы $\mathbf{T}$, удовлетворяющей условию (9). Таким образом, трудоемкости вычисления значений функций $\psi(\mathbf{u})$ и $\varphi_E(\mathbf{u})$ оказываются соизмеримыми.

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

Если $t^* < t_{\max} < \infty$, $u_i(t^*) = 0,\ i \in \{1,...,m_1\}$ (точка $\mathbf{u}(t^*)$ лежит на границе множества $\{\mathbf{u} : u_i \le 0,\ i = 1,...,m_1\}$), то субградиент функции $\gamma_E(\mathbf{u})$ в точке $\mathbf{u}(t^*)$ определяется в соответствии с теоремой 5.

В случае $t^* = t_{\max} < \infty$ возможные решения $\mathbf{x}(\mathbf{u}(t^*))$ задачи (3) описываются теоремой 1. При формулировке следующей теоремы используются обозначения пункта 2.

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

$$-\langle \mathbf{y}, \mathbf{y} \rangle + \langle \mathbf{T}' \mathbf{l^0}, \mathbf{y} \rangle + c^0 = E \tag{28}$$

и соотношениям (12) теоремы 1 ($y_i^* = y_i(t_{\max}),\ i \notin I^*$). При этом вектор $\mathbf{x}^* = \mathbf{T}\mathbf{y}^*$ – решение задачи (3) при $\mathbf{u} = \mathbf{u}(t^*)$, а вектор $\left(K_i(\mathbf{x}^*)\right)_{i=1}^N$ – субградиент функции $\gamma_E(\mathbf{u})$ в точке $\mathbf{u}(t^*)$. Здесь матрица $\mathbf{T}$ удовлетворяет условию (9).

Доказательство. Из теоремы 5 следует, что если для субградиента $\mathbf{g}_\psi$ функции ψ в точке $\mathbf{u}(t^*)$ выполняется

$$\psi(\mathbf{u}(t^*)) - E = (\mathbf{g}_\psi, \mathbf{u}(t^*) - \mathbf{u}^0), \tag{29}$$

то вектор $\mathbf{g}_\psi$ – 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. $\mathbf{g}_\psi = K(\mathbf{x}^*) = \left(K_i(\mathbf{x}^*)\right)_{i=1}^N$. Решения $\mathbf{x}^*$ в соответствии с теоремой 1 могут быть представлены в виде $\mathbf{x}^* = \mathbf{T}\mathbf{y}^*$, где матрица $\mathbf{T}$ и вектор $\mathbf{y}^*$ удовлетворяют условиям теоремы 1.

Пусть $\mathbf{x}^*$ – решение задачи (3) при $\mathbf{u} = \mathbf{u}(t^*)$, для которого выполняется условие (29). Нетрудно видеть, что $\langle \mathbf{g}_\psi, \mathbf{u}(t^*) - \mathbf{u}^0 \rangle = \langle \mathbf{K}(\mathbf{x}^*), \mathbf{u}(t^*) - \mathbf{u}^0 \rangle =$ $= L(\mathbf{x}^*, \mathbf{u}(t^*)) - L(\mathbf{x}^*, \mathbf{u}^0) = \psi(\mathbf{u}(t^*)) - L(\mathbf{x}^*, \mathbf{u}^0)$. Откуда следует $L(\mathbf{x}^*, \mathbf{u}^0) = E$. Используя обозначения пункта 2, имеем $L(x^*, u^0) = \langle \mathbf{A}(\mathbf{u}^0)\mathbf{x}^*, \mathbf{x}^* \rangle +$ $+ \langle \mathbf{l}(\mathbf{u}^0), \mathbf{x}^* \rangle + c(\mathbf{u}^0) = \langle \mathbf{A}^0\mathbf{x}^*, \mathbf{x}^* \rangle + \langle \mathbf{l}^0, \mathbf{x}^* \rangle + c^0$. Делая замену $\mathbf{x}^* = \mathbf{T}\mathbf{y}^*$ и учитывая, что матрица $\mathbf{T}$ удовлетворяет условию (9), получаем $L(\mathbf{x}^*, \mathbf{u}^0) =$ $= -\langle \mathbf{y}^*, \mathbf{y}^* \rangle + \langle \mathbf{T}'\mathbf{l^0}, \mathbf{y}^* \rangle + c^0$. Отсюда получаем, что если условие (28) выполняется, то вектор $\mathbf{g}_\psi$ – субградиент функции γ_E в точке $\mathbf{u} = \mathbf{u}(t^*)$.

Уравнение (28) при дополнительных соотношениях (12) теоремы 1 разрешимо не при любых значениях параметра E. В частности, если $E > \psi(\mathbf{u}^0)$, то уравнение (28) 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 (28) решений не имеет, то найдется точка $t' < t_{\max}$, в которой будет выполняться условие (24). Последнее противоречит предположению о том, что $t^* = t_{\max}$. Теорема доказана. ■

Выводы. В работе для вычисления оценки оптимального значения за-дачи 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. Показано, что 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 будут полезны при разработке эффективных методов решения оценочных задач квадратичной оптимизации.

Автор выражает глубокую благодарность Березовскому О.А. за ценные замечания и советы по данной работе.

Ю.П. Лаптін

КОНІЧНА РЕГУЛЯРІЗАЦІЯ У ЗАДАЧАХ КВАДРАТИЧНОЇ ОПТИМІЗАЦІЇ

Розглядаються питання обчислення оцінок оптимальних значень неопуклих задач квадратичної оптимізації на основі лагранжевих релаксацій вихідної задачі. На границі допустимої області оціночної задачі функції задачі можуть бути розривні, погано обумовлені, що ускладнює розробку обчислювальних алгоритмів. У роботі наводяться нові підходи подолання зазначених проблем, засновані на використанні конічних регуляризації опуклих задач оптимізації.

Yu.P. Laptin

CONIC REGULARIZATION IN QUADRATIC OPTIMIZATION PROBLEMS

Calculation of estimates for optimal values for non-convex quadratic optimization problems on the base of Lagrange relaxation of the original problem is considered. At the boundary of the feasible set of the estimation problem the used functions can be discontinuous or poorly conditioned that complicates the development of numerical algorithms. The paper presents a new approach to overcome these problems, based on the use of conic regularizations of convex optimization problems.

1. *Шор Н.З., Стеценко С.И.* Квадратичные экстремальные задачи и недифференцируемая оптимизация. Киев: Наукова думка, 1989. 204 с.
2. *Березовский О.А., Стецюк П.И.* Об одном способе нахождения двойственных квадратичных оценок Шора. *Кибернетика и системный анализ*. 2008. № 2. С. 89 – 99.
3. *Boyd S., Vandenberghe L.* Semidefinite programming relaxations of non-convex problems in control and combinatorial optimization. *Communications, Computation, Control, and Signal Processing. – Springer US*, 1997. P. 279 – 287.
4. *Поляк Б.Т., Щербаков П.С.* Рандомизированный метод решения задач полуопределенного программирования. *Стохастическая оптимизация в информатике*. – Издательство Санкт-Петербургского государственного университета (Санкт-Петербург), 2006. Том 2. № 1-1. С. 38 – 70.
5. *Лаптин Ю.П., Бардадым Т.А.* Некоторые подходы к регуляризации нелинейных задач оптимизации. *Проблемы управления и информатики*. 2011. № 3. С. 57 – 68.
6. *Беллман Р.* Введение в теорию матриц. М.: Наука, 1976. С. 352.
7. *Михалевич В.С., Трубин В.А., Шор Н.З.* Оптимизационные задачи производственно-транспортного планирования. М.: Наука, 1986. С. 260.

Получено 23.09.2016

Об авторе:

Лаптин Юрий Петрович,

доктор физико математических наук, старший научный сотрудник
Института кибернетики имени В.М. Глушкова НАН Украины.