ТЕОРІЯ ОПТИМАЛЬНИХ РІШЕНЬ

УДК 519.8

О.А. БЕРЕЗОВСКИЙ, Г.А. ШУЛИНОК

АНАЛИТИЧЕСКОЕ ВЫРАЖЕНИЕ ДВОЙСТВЕННОЙ ОЦЕНКИ ДЛЯ ЗАДАЧИ РАЗМЕЩЕНИЯ ТОЧЕК В ШАРЕ

Получено аналитическое выражение для двойственной оценки оптимального значения целевой функции квадратичной постановки задачи размещения точек в шаре. Этот результат справедлив также для оценки, получаемой в результате применения к данной постановке задачи SDP-релаксации.

Задачи размещения и упаковки возникают при оптимизации хранения и транспортировки товаров, при исследовании кристаллических или пористых структур, при проектировании и компоновке различных технологических объектов и систем, в теории надежности системы сигналов, при анализе структуры нефтеносных пород, при моделировании процессов в активной зоне ядерного реактора и т. д. Предметом исследования данной работы является задача размещения m точек в шаре единичного радиуса так, чтобы минимальное расстояние между ними (равное x_0^2) было максимальным [1, 2]:

$$f^* = \max x_0^2, \tag{1}$$

$$x_0^2 - \sum_{k=1}^{n} (x_{ik} - x_{jk})^2 \le 0,\ 1 \le i < j \le m, \tag{2}$$

$$\sum_{k=1}^{n} x_{ik}^2 - 1 \le 0,\ i = \overline{1,m}. \tag{3}$$

Поскольку данная задача относится к NP-трудным, становится актуальной оценка глобального оптимума, например, путем вычисления двойственной оценки [3, 4]

$$\psi^* = \inf_{\substack{\lambda_{\max}(A(u,v)) \le 0 \\ u,v \le 0}} \Big(\psi(u,v) = \sup_{x \in R^{nm+1}} L(x,u,v) \Big) \ge f^*, \tag{4}$$

где $L(x,u,v) = x^T A(u,v) x + c(u)$ – функция Лагранжа для задачи (1) – (3); u – вектор двойственных переменных, компонента u_{ij} которого соответствует ограничению вида (2) для i-ой и j-ой точек, $1 \le i < j \le m;$ v – вектор двойственных переменных для ограничений (3); эффективное множество $\bar{D}$ двойственных переменных для однородной квадратичной задачи $\psi(u,v) = \sup_x L(x,u,v)$ записано в виде неположительности максимального собственного числа матрицы $A(u,v)$ $\lambda_{\max}(\mathrm{A}(u,v)) \le 0.$ Свободный член функции Лагранжа $c(u) = -\sum_{i=1}^{m} v_i,$ а матрица $A(u,v)$ представляет собой блочную диагональную матрицу, первый блок которой состоит из одного элемента $1 + \sum_{\substack{i,j=1 \\ j>i}}^{m} u_{ij}$ (соответствует переменной x_0), а остальные n блоков (для каждой k-ой координаты) одинаковы

$$\begin{array}{c} \\ x_{1k} \\ x_{2k} \\ \\ x_{mk} \end{array} \begin{array}{c} \begin{array}{cccc} x_{1k} & x_{2k} & & x_{mk} \end{array} \\ \begin{pmatrix} -\sum_{j>1}^{m} u_{1j} + v_1 & u_{12} & \dots & u_{1m} \\ u_{12} & -u_{12} - \sum_{j>2}^{m} u_{2j} + v_2 & \dots & u_{2m} \\ \dots & \dots & \dots & \dots \\ u_{1m} & u_{2m} & \dots & -\sum_{j=1}^{m-1} u_{jm} + v_m \end{pmatrix} \end{array}. \qquad (5)$$

Вследствие однородности квадратичной задачи (1) – (3), решение внутренней задачи нахождения функции $\psi(u,v) = \sup_x L(x,u,v)$ в (4) при $u \in \bar{D}$ тривиально:

$$\psi(u) = \begin{cases} -\infty, & if\ \lambda_{\max}(A(u)) > 0 \\ c(u), & if\ \lambda_{\max}(A(u)) \le 0 \end{cases}$$

и вычисление двойственной оценки (4) для рассматриваемой задачи сводится к решению задачи выпуклой недифференцируемой оптимизации

$$\psi^* = \min_{u \in R^n}(-\sum_{i=1}^{m} v_i), \qquad (6)$$

$$\lambda_{\max}(A(u,v)) \le 0, \quad u,v \le 0, \qquad (7)$$

$$u,v \le 0. \qquad (8)$$

Утверждение 1. При $N > m + \psi^*$ точки минимума задачи (6) – (8) и задачи

$$\min_{\substack{u,v \\ u \le 0}} \left(-\sum_{i=1}^{m} v_i + N \max\left(0; \lambda_{\max}(A(u,v)); v_i, i = \overline{1,m}\right) \right) \qquad (9)$$

совпадают.

Доказательство. Для доказательства нам понадобится результат, сформулированный далее в теореме 1 (стр. 25, [5]). Рассмотрим задачу

$$\min f_0(y), \tag{10}$$

$$f_i(y) \le 0, , \ i = \overline{1,k}, \tag{11}$$

$$y \in M, \tag{12}$$

где $f_i(y)$, $i = \overline{0,k}$, – непрерывные функции, $M \subseteq R^n$ – некоторое множество. Введем семейство задач, зависящее от вектора параметров $z \in R^k$,

$$V(z) = \inf\{ f_0(y) : \ f_i(y) \le z_i, i = \overline{1,k}; y \in M \}.$$

Если воспользоваться для учета ограничений (11) негладкой штрафной функцией в форме функции максимума

$$\Phi_N(y) = f_0(y) + N \max\{0, f_1(y), \ldots, f_k(y)\}, \tag{13}$$

то справедлива

Теорема 1 (стр. 25, [5]). Пусть $\inf\limits_{t>0} \dfrac{V(te) - V(0)}{t} = -L > -\infty$, где e – k-мерный вектор, все компоненты которого равны единице, $t \in R^1$, и $N > L$. Тогда точки минимума исходной задачи $V(0)$ и задачи $\inf\limits_{y \in M} \Phi_N(y)$, где $\Phi_N(y)$ определяется по формуле (13), совпадают.

Запишем функцию $V(te)$ для задачи (10) – (12)

$$V(te) = \inf\left\{ -\sum_{i=1}^{m} v_i : \lambda_{\max}(A(u,v)) \le t, v_i \le t, i = \overline{1,m}, u \in M \right\} =$$

$$= \inf\left\{ -\sum_{i=1}^{m} v_i : \lambda_{\max}(A(u,v) - tI) \le t, v_i - t \le 0, i = \overline{1,m}, u \in M \right\},$$

где $M = \{u : u \le 0\}$. Исследуем ее, разбив интервал $(0, +\infty)$ возможных значений параметра t на два случая: $(0,1)$ и $[1, +\infty)$.

1) в случае, когда $t \ge 1$, первыми строкой и столбцом матрицы $(A(u,v) - tI)$ можно пренебречь, так как $A_{11} = 1 + \sum\limits_{\substack{i,j=1 \\ j>i}}^{m} u_{ij} - t < 0 \quad \forall u \in M.$

Из неположительности других диагональных элементов матрицы $(A(u,v) - tI)$ (необходимое условие неположительной определенности матрицы) $A_{ii}^{(k)} =$ $= -\sum\limits_{j=1}^{j<i} u_{ji} - \sum\limits_{j>i}^{m} u_{ij} + v_i - t \le 0$ (под $A_{ij}^{(k)}$ будем понимать (i,j)-элемент k-го блока матрицы $(A(u,v) - tI)$) следует, что $v \le te_v$. Но точка $(\overline{u}, \overline{v}) = (0_u, te_v)$

с максимально возможными значениями элементов вектора v является допустимой и, следовательно, $V(te) = -\sum_{i=1}^{m} \bar{v}_i = -mt.$ Тогда оценка штрафного множителя

$$\inf_{t\ge 1} \frac{V(te)-V(0)}{t} = \inf_{t\ge 1}\left(\frac{-mt-V(0)}{t}\right) = -m - V(0) > -\infty;$$

2) случай $0 < t < 1.$ Сделаем замену переменных $\tilde{u}_{ij} = \dfrac{u_{ij}}{1-t},\ \ 1 \le i < j \le m,$ $\tilde{v}_i = \dfrac{v_i - t}{1-t},\ \ i = \overline{1,m}.$ Тогда элементы матрицы $(A(u,v) - tI)$ примут вид

$$A_{11} = 1 + \sum_{\substack{i,j=1\\ j>i}}^{m} u_{ij} - t = (1-t)\left(1 + \sum_{\substack{i,j=1\\ j>i}}^{m} \frac{u_{ij}}{1-t}\right) = (1-t)\left(1 + \sum_{\substack{i,j=1\\ j>i}}^{m} \tilde{u}_{ij}\right),$$

$$A_{ii}^{(k)} = -\sum_{j=1}^{j<i} u_{ji} - \sum_{j>i}^{m} u_{ij} + v_i - t = (1-t)\left(-\sum_{j=1}^{j<i} \tilde{u}_{ji} - \sum_{j>i}^{m} \tilde{u}_{ij}\tilde{v}_i\right) +$$

$$+(1-t)\left(\frac{v-t}{1-t}\right) = (1-t)\left(-\sum_{j=1}^{j<i} \tilde{u}_{ji} - \sum_{j>i}^{m} \tilde{u}_{ij} + \tilde{v}_i\right),$$

$$A_{ij}^{(k)} = u_{ij} = (1-t)\tilde{u}_{ij}.$$

В новых переменных функция $V(te)$ примет вид

$$V(te) = \inf\left\{-\sum_{i=1}^{m}\big((1-t)\tilde{v}_i + t\big) : (1-t)\lambda_{\max}(A(\tilde{u},\tilde{v})) \le 0, (1-t)\tilde{v} \le 0, \tilde{u} \in \tilde{M}\right\},$$

где $\tilde{M} = \{\tilde{u} : (1-t)\tilde{u} \le 0\}.$ Учитывая, что рассматривается случай $1 - t > 0,$

$$V(te) = -\sum_{i=1}^{m} t + (1-t)\inf\left\{-\sum_{i=1}^{m} \tilde{v}_{ik} : \lambda_{\max}(A(\tilde{u},\tilde{v})) \le 0, (\tilde{u},\tilde{v}) \in M\right\} =$$

$$= -mt + (1-t)V(0).$$

Тогда

$$\inf_{0<t<1} \frac{V(te)-V(0)}{t} = \inf_{0<t<1} \frac{-mt + (1-t)V(0) - V(0)}{t} = -m - V(0) > -\infty.$$

Таким образом,

$$\inf_{t>0} \frac{V(te)-V(0)}{t} = \min\left(\inf_{0<t<1} \frac{V(te)-V(0)}{t}, \inf_{t\ge 1} \frac{V(te)-V(0)}{t}\right) = -m - \psi^* = -L > -\infty. .$$

С учетом теоремы 1 доказательство утверждения завершено.

Сформулируем основной результат данной работы.

Утверждение 2. Двойственная оценка (4) задачи (1) – (3) равна

$$\psi^* = \frac{2m}{(m-1)}.$$

Доказательство. Рассмотрим задачу (9), к которой, как показано выше, сводится нахождение двойственной оценки задачи (1) – (3). Пусть имеем точку (u^*, v^*), причем $u_{ij}^* < 0$, $1 \le i < j \le m$. Тогда для того, чтобы доказать, что (u^*, v^*) является решением задачи (9), достаточно показать, что множество субградиентов целевой функции задачи (9)

$$\Psi(u,v) = -\sum_{i=1}^{m} v_i + N \max\left(0; \lambda_{\max}(A(u,v)); v_i, i = \overline{1,m}\right)$$

в этой точке содержит нулевой вектор (в случае $u^* < 0$ ограничением можно пренебречь и задачу (9) можно рассматривать как безусловную задачу минимизации выпуклой недифференцируемой функции). Покажем, что это условие выполняется для точки (u^*, v^*), определенной следующим образом:

$$u_{ij}^* = -\frac{2}{m(m-1)},\ 1 \le i < j \le m,\ \ v_i^* = -\frac{2}{(m-1)},\ i = \overline{1,m}.$$

Нетрудно видеть, что при этих значениях двойственных переменных $A(u^*, v^*)$ является отрицательно полуопределенной матрицей ранга n, в которой ненулевые собственные числа равны $\lambda_k = -\frac{2}{(m-1)}$, $k = \overline{1,n}$ (по одному для каждого блока), а соответствующие им собственные вектора определяются $\xi_k = (0_1 \eta_1 \eta_2 \dots \eta_n)^T \in R^{nm+1}$, где вектор-строки $\eta_l = 0_m \in R^m$ для всех $l \ne k$, $l = \overline{1,n}$ – нулевые, а при $l = k$ – $\eta_k = \frac{1}{\sqrt{m}} e_m^T \in R^m$ (0_m и e_m – нулевой и единичный m-мерные вектор-строки). Для наглядности приведем соответствующий данной точке (u^*, v^*) вид функции Лагранжа постановки (1) – (3)

$$L(z, x, u^*, v^*) = -\frac{2}{m(m-1)} \sum_{k=1}^{n} \left(\sum_{i=1}^{m} x_{ik} \right)^2 + \frac{2m}{(m-1)}.$$

Покажем, что в выбранной точке (u^*, v^*) существует нулевой субградиент целевой функции $\Psi(u,v)$ задачи (9) – $0 \in \partial\Psi(u^*, v^*) = \mathrm{conv}\{\Psi'(u^*, v^*)\}$. Координаты субградиентов $\Psi'(u^*, v^*)$ определяются по формулам:

$$\Psi'_{u_{ij}}(u^*, v^*) = N\alpha\left(\lambda_{\max}(A(u^*, v^*))\right)'_{u_{ij}},\ \ \Psi'_{v_i}(u^*, v^*) = -1 + N\alpha\left(\lambda_{\max}(A(u^*, v^*))\right)'_{v_i},$$

где коэффициент $\alpha,\ 0 \le \alpha \le 1,$ введен для учета того, что

$$\left(\max\{0, \lambda_{\max}(A(u^*, v^*))\}\right)' \in conv\{0, \left(\lambda_{\max}(A(u^*, v^*))\right)'\},$$

а субградиент функции максимального собственного числа матрицы $A(u,v)$ в точке (u^*, v^*) $\left(\lambda_{\max}(A(u^*, v^*))\right)' \in \mathrm{conv}(p : p \in \partial\lambda_{\max}(A(u^*, v^*)))$, $p \in R^{m+m(m-1)/2}$, где компоненты вектора p определяются из равенств

$$p_{u_{ij}} = s_0^2 + \sum_{k=1}^{n}(-s_{x_{ik}}^2 + 2s_{x_{ik}}s_{x_{jk}} - s_{x_{jk}}^2),\ \ p_{v_i} = \sum_{k=1}^{m} s_{x_{ik}}^2. \tag{14}$$

Здесь $s \in \{\xi_l, l = \overline{n+1, mn+1}\} \in R^{nm+1}$ – любой вектор из набора собственных векторов матрицы $A(u^*, v^*)$, соответствующих ее нулевым собственным числам, а $s_{x_{ik}}$ – его координата, соответствующая переменной x_{ik}. Другими словами, с учетом того, что собственное число равно нулю, s представляет собой произвольный нормированный вектор, принадлежащий ортогональному дополнению к подпространству $\mathrm{conv}(\xi_l,\ l = \overline{1,n})$, определенному выше.

Определим вектор $s^{(i,j)} \in R^{nm+1}$ следующим образом: $(nm+1-2n)$ его компонент равны нулю, и только компоненты, соответствующие x_{ik} и x_{jk} равны $\dfrac{1}{\sqrt{2n}}$ и $-\dfrac{1}{\sqrt{2n}}$ соответственно, $k = \overline{1,n}$. Рассмотрим набор векторов, состоящий из $\dfrac{m(m-1)}{2}$ векторов $\mathrm{S} = \{s^{(i,j)}, 1 \le i < j \le m\}$ и вектора $s^0 =$ $= (1, 0_{nm})^T$ (ненулевая координата соответствует переменной x_0). Для этого набора легко проверить, что

1) все векторы нормированы ($\left\|s^0\right\| = 1$, $\left\|s^{(i,j)}\right\| = 1\ \ \forall s^{(i,j)} \in S$) и принадлежат ортогональному дополнению к $\mathrm{conv}(\xi_l, l = \overline{1,n})$ $((s^{(i,j)}, \xi_l) = 0\ \ \ \forall s^{(i,j)} \in S$, $(s^0, \xi_l) = 0$, $l = \overline{1,n}$);

2) вектор $p^{(i,j)}$, вычисленный для $s^{(i,j)} \in S$ по формулам (14), имеет все нулевые компоненты, кроме $p_{u_{ij}}^{(i,j)} = \sum_{k=1}^{n}\left(-\dfrac{2}{n}\right) = -2$, $p_{u_{li}}^{(i,j)} = \sum_{k=1}^{n}\left(-\dfrac{1}{2n}\right) = -\dfrac{1}{2}$,

$1 \le l < i, \quad p^{(i,j)}_{u_{il}} = -\frac{1}{2}, \quad i < l \le m, \quad l \ne j, \quad p^{(i,j)}_{u_{lj}} = -\frac{1}{2}, \quad 1 \le l < j, \quad l \ne i,$ $p^{(i,j)}_{u_{jl}} = -\frac{1}{2}, \quad j < l \le n, \quad p^{(i,j)}_{v_i} = p^{(i,j)}_{v_j} = \sum_{k=1}^{n} \frac{1}{2n} = \frac{1}{2};$ компоненты вектора p^0, вычисленного для s^0 по формулам (14), равны $p^0_u = e_{m(m-1)/2}$, $p^0_v = 0_m$.

Рассмотрим линейную комбинацию $p = \beta p^0 + \sum_{p^{(i,j)} \in S} p^{(i,j)}$. Поскольку

$$\sum_{p^{(i,j)} \in S} p^{(i,j)}_u = -\left(2 + \frac{m-2}{2}\right) e_{m(m-1)/2} \quad \text{и} \quad \sum_{p^{(i,j)} \in S} p^{(i,j)}_v = \frac{m-1}{2} e_m,$$

то при $\beta = 2 + \frac{m-2}{2}$ справедливо

$$\Psi'_{u_{ij}}(u^*, v^*) = N\alpha p_{u_{ij}} = N\alpha(\beta p^0_{u_{ij}} + \sum_{p^{(i,j)} \in S} p^{(i,j)}_{u_{ij}}) = 0 \quad \forall \alpha,$$

$$\Psi'_{v_i}(u^*, v^*) = -1 + N\alpha p_{v_i} = 0 \text{ при } \alpha = \left(\frac{m-1}{2} N\right)^{-1}.$$

Итак, мы показали, что в точке $(u^*, v^*) = (-\frac{2}{m(m-1)} e_{m(m-1)/2}, -\frac{2}{(m-1)} e_m)$, которая является внутренней точкой допустимого множества задачи (9) ($u^* < 0$), справедливо $0 \in \partial\Psi(u^*, v^*) = \operatorname{conv}\{\Psi'(u^*, v^*)\}$, т. е. она является точкой минимума выпуклой задачи (9); значение функции $\Psi(u, v)$ в этой точке равно

$$\Psi(u^*, v^*) = -\sum_{i=1}^{m} v^*_i = \frac{2m}{(m-1)}.$$

Доказательство завершено.

Следствие 1. Квадрат расстояния между точками в задаче о размещении m точек в шаре единичного радиуса в n-мерном пространстве не превышает $\frac{2m}{(m-1)}$.

Следствие 2. Оценка, полученная с помощью SDP-релаксации задачи (1) – (3), равна $\frac{2m}{(m-1)}$.

О.А. Березовський, Г.О. Шулінок

АНАЛІТИЧНИЙ ВИРАЗ ДВОЇСТОЇ ОЦІНКИ ДЛЯ ЗАДАЧІ РОЗМІЩЕННЯ ТОЧОК У КУЛІ

Отримано аналітичний вираз для двоїстої оцінки оптимального значення цільової функції квадратичної постановки задачі розміщення точок у кулі. Цей результат справедливий також для оцінки, що отримується в результаті застосування до даної постановки задачі SDP-релаксації.

O.A. Berezovskyi, G.A. Shulinok

ANALYTICAL EXPRESSION OF THE DUAL BOUND FOR THE POINT-PACKING PROBLEM IN THE BALL

We obtain the analytical expression for the dual bound of the optimal value for the quadratic formulation of the point-packing problem in the ball. This result is valid also for the bounds, obtained by invoking SDP-relaxation to this formulation of the problem.

1. *Conway J.H. and Sloane N.J.A.* Sphere Packings, Lattices and Groups (Second Edition). – NY: Springer-Verlag, 1993. – 679 p.
2. *Anstreicher K.* On convex relaxations for quadratically constrained quadratic programming // Mathematical programming. – 2012. – N 136 (2). – P. 233 – 251.
3. *Шор Н.З., Стеценко С.И.* Квадратичные экстремальные задачи и недифференцируемая оптимизация. – Киев: Наук. думка, 1989. – 208 с.
4. *Shor N.Z.* Nondifferentiable Optimization and Polynomial Problems. – Dordrecht, Kluwer, 1998. – 394 p.
5. *Пшеничный Б.Н.* Метод линеаризации. – М.: Наука, 1983. – 136 с.

Получено 24.03.2014