УДК 519.8

О.А. БЕРЕЗОВСКИЙ, К.А. ЖЕРЕБ

ДВОЙСТВЕННЫЕ ОЦЕНКИ ДЛЯ ЗАДАЧИ О МАКСИМАЛЬНОМ *K*-КЛАБЕ*

Рассмотрены возможности формирования квадратичных постановок задачи о максимальном k-клабе и, соответственно, нахождения верхних двойственных оценок, получаемых с помощью техники Н.З. Шора. На примере задачи о максимальном 2-клабе показано, что неоднозначность построения соответствующих квадратичных задач (в том числе с учетом добавления функционально избыточных ограничений) предоставляет значительные возможности для уточнения двойственных оценок.

При исследовании сетевых структур (в социологии, медицине, биологии, транспорте и т.д.) для анализа связных подмножеств удобно строить графовые модели, использующие такие понятия как k-клика, k-клаб, k-плекс и др. [1]. Эти понятия введены для характеристики структурных свойств выделяемого подмножества (определяют критерий связности его элементов) и, по сути, могут рассматриваться как модификации достаточно известной кликовой модели [2], расширяя (с точки зрения "ослабления" связей, при $k \geq 2$) возможности при моделировании реальных ситуаций в сетях [3]. Возникающие при этом экстремальные задачи на графах как благодаря растущему к ним практическому интересу (сетевые задачи приобретают все большую значимость в обществе), так и в силу своей NP-полноты привлекают внимание исследователей, исповедывающих различные математические школы. В отличие от большинства авторов, которые при рассмотрении данных проблем не выходили за рамки линейных моделей, мы предлагаем применять квадратичные модели, что позволит воспользоваться подходом академика Н.З. Шора для получения двойственных лагранжевых оценок с использованием функционально избыточных ограничений [4]. Применение данного подхода приводит к относительно

* Работа выполнена при частичной финансовой поддержке гранта UKM2-2812-KV-06 (CRDF Cooperative Grants Program).

неплохим, а в ряде случаев и к точным оценкам решений задач [4]. Причем, в первом случае всегда остается надежда, что расширение использованной квадратичной постановки за счет введения (добавления) новых избыточных ограничений позволит улучшить полученную для нее оценку и привести к точному значению по функционалу. Следует отметить, что этот подход ни в коем случае не сужает используемый для исследований математический аппарат – имеется в виду, что на базе расширенных квадратичных моделей можно строить и линейные модели.

В данной работе рассмотрим возможности применения техники двойственных квадратичных оценок на примере одной из "простейших" из вышеупомянутых кластерных задач – задачи о максимальном 2-клабе. Для начала сформулируем задачу о максимальном k-клабе в общем виде.

Пусть задан неориентированный граф $G(V,E)$ с множеством вершин $V=\{1,\dots,n\}$ и множеством ребер $E=\{(i,j): i\in V, j\in V\}$. Под k-клабом графа $G(V,E)$ подразумевают подмножество вершин графа $G(V,E)$, из каждой вершины которого можно попасть в другую не более чем через $(k-1)$ точку этого подмножества, соединенных ребрами из множества E (при $k=1$ получаем определение клики графа $G(V,E)$). Задача о максимальном k-клабе заключается в выделении в графе $G(V,E)$ k-клаба максимальной мощности (т. е. содержащего максимальное количество вершин). Эту задачу можно сформулировать в виде следующей задачи дискретного программирования [3]:

$$\overline{w}_k(G) = \max \sum_{i\in V} x_i \tag{1}$$

при ограничениях

$$x_i + x_j \le 1 + \sum_{l:P_{ij}^l \in P_{ij}} y_{ij}^l \quad \forall (i,j)\notin E, \tag{2}$$

$$x_p \ge y_{ij}^l \quad \forall p \in V(P_{ij}^l),\ \forall l : P_{ij}^l \in P_{ij},\ \forall (i,j)\notin E, \tag{3}$$

$$x_i \in \{0,1\} \quad i\in V, \tag{4}$$

$$y_{ij}^l \in \{0,1\} \quad \forall P_{ij}^l \in P_{ij},\ \forall (i,j)\notin E. \tag{5}$$

Здесь каждой вершине графа соответствует булевая переменная x_i, которая принимает значение 1, если вершина i принадлежит k-клабу, и 0, в противном случае; $P_{ij}=\{P_{ij}^l\}_l$ – множество всех путей между вершинами i и j в графе G, длина которых не более k, а P_{ij}^l – элемент этого множества (путь с номером l), которому ставится в соответствие булевая переменная y_{ij}^l – равна 1 только в том случае, если все вершины x_p данного пути равны 1 (т.е. принадлежат k-клабу). Смысл ограничений (2) – (5) состоит в следующем: если две вершины i

и j принадлежат k-клабу, то $1 \le \sum_{l: P_{ij}^l \in P_{ij}} y_{ij}^l$ (из условия (2)), т. е. существует хотя бы один путь между данными вершинами (длина которого не более k), все вершины которого также принадлежат k-клабу ($x_p \ge y_{ij}^l$). Если же такого пути не существует, то $\sum_{l: P_{ij}^l \in P_{ij}} y_{ij}^l = 0$ и согласно (2) $x_i + x_j \le 1$, т. е. хотя бы одна вершина не принадлежит k-клабу.

Построение квадратичной постановки рассматриваемой задачи далеко неоднозначно и, на сегодняшний день, не имеет общего условия предпочтительности той или иной из них (кроме, конечно, критерия накопления ограничений, который, однако, может "раздуть" размеры даже небольшой задачи до неприемлемой с вычислительной точки зрения размерности). Понятно, что самая простая квадратичная постановка с непрерывными переменными получается из задачи (1) – (5) после выражения условий булевости переменных квадратичными равенствами $x_i^2 - x_i = 0$, $i \in V$, и $(y_{ij}^l)^2 - y_{ij}^l = 0$, $\forall P_{ij}^l \in P_{ij}$, $\forall (i,j) \notin E$. Если воспользоваться другими приемами, например, результатами различных допустимых перемножений линейных выражений, которые либо явно присутствуют в условиях задачи (1) – (5), либо являются их следствиями (например, $0 \le x_i \le 1$ или $0 \le y_{ij}^l \le 1$), то соответствующие ограничения можно заменить квадратичными и получить ряд других квадратичных формулировок задачи о максимальном k-клабе даже без наличия каких-либо функционально избыточных ограничений. В случае если в постановку включать группу или даже все из таких возможных ограничений-следствий (при этом появятся и функционально избыточные ограничения), то двойственная квадратичная оценка как минимум не ухудшится, а в большинстве случаев улучшится за счет большей "мобильности" матрицы функции Лагранжа (в смысле увеличения области положительной определенности квадратичной формы функции Лагранжа).

Сформулируем одну из таких квадратичных постановок задачи о максимальном k-клабе (без функционально избыточных ограничений):

$$\overline{w}_k(G) = \max \sum_{i \in V} x_i \tag{6}$$

при ограничениях

$$x_i x_j \le x_i \sum_{l: P_{ij}^l \in P_{ij}} y_{ij}^l \quad \forall (i,j) \notin E, \tag{7}$$

$$x_p \ge y_{ij}^l \quad \forall p \in V(P_{ij}^l), \ \forall l : P_{ij}^l \in P_{ij}, \ \forall (i,j) \notin E, \tag{8}$$

$$x_i^2 - x_i = 0, \ i \in V, \tag{9}$$

$$(y_{ij}^l)^2 - y_{ij}^l = 0\,,\ \forall P_{ij}^l \in P_{ij}\,,\ \forall (i,j) \notin E\,. \tag{10}$$

В данном случае ограничение (7) получено путем умножения обеих частей неравенства (2) на x_i с учетом равенства $x_i^2 - x_i = 0$ (заметим, что для каждой пары $(i,j) \notin E$ получаем два неравенства – как при умножении на x_i, так и на x_j). При этом смысл ограничения (7) остался тем же, что и у ограничения (2).

При выборе приведенной постановки мы руководствовались следующими соображениями. Во-первых, стремлением увеличить количество двойственных переменных в матрице функции Лагранжа, которые могли бы помочь расширить ее область положительной определенности (например, ограничения (7) это позволяют, в то время как замена (8) на $x_p y_{ij}^l \geq y_{ij}^l$ ничего не дает). Во-вторых, нежеланием значительно увеличивать количество ограничений (этим объясняется отсутствие в постановке функционально избыточных ограничений, а также отказ от полиномиальной постановки задачи о максимальном k-клабе, которая с одной стороны содержит только переменные x_i, $i \in V$, но влечет за собой последующее значительное увеличение переменных при сведении к квадратичному виду). И, в-третьих, тем, что ограничение (7) активно (в том смысле, что может превращаться в равенство) на большей области определения прямых переменных задачи, чем другие возможные ограничения-следствия (например,

$$x_i x_j \leq \sum_{l: P_{ij}^l \in P_{ij}} y_{ij}^l\,,\ x_i x_j \leq \left(\sum_{l: P_{ij}^l \in P_{ij}} y_{ij}^l \right)^2,\ \left(x_i + x_j\right)^2 \leq 1 + \left(\sum_{l: P_{ij}^l \in P_{ij}} y_{ij}^l \right)^2$$

и т.п.), что также предполагает большую гибкость при работе с матрицей функции Лагранжа, а значит и вероятность получения более точных оценок.

В частном случае, когда $k = 2$ (заметим, что это справедливо и при $k = 3$), задача о максимальном k-клабе значительно упрощается – переменные y_{ij}^l исключаются (каждому пути l соответствует только одна вершина и переменные x_p, y_{ij}^l принимают один и тот же смысл) и размерность задачи становится равной $|V|$. Для данной задачи выпишем два варианта квадратичной постановки из упоминавшихся выше:

– вариант 1 (простейший)

$$\overline{w}_2(G) = \max \sum_{i \in V} x_i \tag{11}$$

при ограничениях

$$x_i + x_j \leq 1 + \sum_{p \in N(i,j)} x_p \quad \forall (i,j) \notin E\,, \tag{12}$$

$$x_i^2 - x_i = 0\,,\ i \in V\,. \tag{13}$$

– вариант 2 (следует из задачи (6) – (10))

$$\overline{w}_2(G) = \max \sum_{i \in V} x_i \tag{14}$$

при ограничениях

$$x_i x_j \le x_i \sum_{p \in N(i,j)} x_p \,,\ x_i x_j \le x_j \sum_{p \in N(i,j)} x_p \quad \forall (i,j) \notin E \,, \tag{15}$$

$$x_i^2 - x_i = 0 \,,\ i \in V \,, \tag{16}$$

где $N(i,j) = N(i) \cap N(j)$ обозначает множество вершин графа, соединенных ребрами как с i-й вершиной ($N(i)$), так и с j-й вершиной ($N(j)$).

На основе приведенных задач (11) – (13) и (14) – (16) проиллюстрируем поведение двойственных оценок Шора на двух специально подобранных тестовых примерах. Обозначим двойственную оценку для задачи (11) – (13) – ψ_1, а для задачи (14) – (16) – ψ_2.

Пример 1. Пусть задан граф, который представляет собой цикл из 9 вершин, дополненный еще одним ребром, соединяющим вершины 2 и 4, т. е. множество ребер равно $E = \{(1,2),(2,3),(3,4),(4,5),(5,6),(6,7),(7,8),(8,9),(9,1),(2,4)\}$. Мощность максимального 2-клаба для этого графа равна $\overline{w}_2(G) = 4$ и достигается в точках (1,1,1,1,0,0,0,0,0) и (0,1,1,1,1,0,0,0,0). Двойственная оценка для задачи (11) – (13) $\psi_1 = 4.5$ не является точной (отметим, что она совпала с линейной верхней оценкой, которая получается в результате релаксации задачи при замене ограничений (13) на $0 \le x_i \le 1$, $i \in V$). Для второго варианта квадратичной постановки оценка улучшилась и, более того, стала точной – $\psi_2 = 4.0$, что показывает преимущество квадратичной постановки (14) – (16). Однако она тоже не всегда будет точной. Покажем это на следующем примере.

Пример 2. Граф представляет собой цикл из 7 вершин. Для этого графа $\overline{w}_2(G) = 3$ и исходная задача имеет семь решений. Оценка $\psi_1 = 3.5$. Во втором случае оценка улучшилась – $\psi_2 = 3.3317667$, но также не является точной. Однако, как уже отмечалось, для таких случаев в технике двойственных квадратичных оценок [4] в запасе имеется достаточно успешный аппарат расширения квадратичных задач за счет функционально избыточных ограничений. Запишем новые ограничения вида

$$x_i x_l + x_j x_l \le x_l \,,\ \ \forall (i,j) \notin E \,,\ N(i,j) = \{\varnothing\} \,, \tag{17}$$

которые строятся домножением линейных ограничений (12) для пар вершин i и j, между которыми нет пути длинной менее трех ребер (т.е. для которых справедливо $\sum_{p \in N(i,j)} x_p = 0$), на переменные x_l. После добавления этих ограничений в задачу (14) – (16) получаем третий вариант квадратичной постановки для задачи о максимальном 2-клабе – задачу (14) – (17), двойственная оценка для которой оказалась точной $\psi_3 = 3.0$.

На предложенных тестовых примерах прослеживаются возможности квадратичных моделей, неоднозначность построения которых (в том числе с использованием функционально избыточных ограничений) предоставляет значительные возможности для уточнения двойственных лагранжевых оценок. Можно сделать вывод, что применение техники двойственных квадратичных оценок Шора [4] для рассматриваемого класса задач (впрочем, как и для всех квадратичных задач) представляет определенный интерес и является достаточно перспективным, однако при этом приобретает значимость решение следующих вопросов: каким образом добавлять функционально избыточные ограничения (по какому правилу), чтобы не вводить слишком много дополнительных переменных, общий критерий возможности достижения точной оценки и т.д.

О.А. Березовський, К.А. Жереб

ДВОЇСТІ ОЦІНКИ ДЛЯ ЗАДАЧІ ПРО МАКСИМАЛЬНИЙ *K*-КЛАБ

Розглянуті можливості формування квадратичних постановок задачі про максимальний *k*-клаб і, відповідно, знаходження верхніх двоїстих оцінок, які можна отримати за допомогою техніки Н.З. Шора. На прикладі задачі про максимальний 2-клаб показано, що неоднозначність побудови відповідних квадратичних задач (у тому числі з урахуванням додавання функціонально надлишкових обмежень) надає значні можливості для уточнення двоїстих оцінок.

O.A. Berezovskyi, K. A. Zhereb

DUAL BOUNDS FOR *K*-CLUB PROBLEM

The opportunities of constraining quadratic formulations of k-club problem and, accordingly, finding of upper dual bounds received by using Shor's engineering are considered. On an example of 2-club problem it's shown, that the ambiguity of construction of the appropriate quadratic problems (including in view of addition of functional superfluous restrictions) gives significant opportunities to improve of dual bounds.

1. *Mokken R.* Cliques, clubs and clans // Quality and Quantity. – 1979. – Vol. 13. – P. 161–173.
2. *Luce R., Perry A.* A method of matrix analysis of group structure // Psychometrika. – 1949. – Vol. 14. – P. 95–116.
3. *Balansundaram B., Butenko S., Trukhanov S.* Novel approaches for analyzing biological networks // J. of Combinatorial Opt. – 2005. – Vol. 10. – P. 23–39.
4. *Shor N.Z.* Nondifferentiable Optimization and Polynomial Problems. – Dordrecht, Kluwer, 1998. – 394 p.

Получено 12.03.2008