Журнал обчислювальної та прикладної математики | 2017, №1(124) | Journal of Numerical & Applied Mathematics

УДК 519.85
MSC 519.85

QUADRATIC PROBLEM FOR MAXIMUM k-PLEX IN UNDIRECTED GRAPH

Petro Stetsyuk[1], Tamara Bardadym[2], Volodymyr Lyashko[3]

[1]Glushkov Institute of Cybernetics of NASU, Kyiv, Ukraine, E-mail: stetsyukp@gmail.com
[2]Glushkov Institute of Cybernetics of NASU, Kyiv, Ukraine, E-mail: tbardadym@gmail.com
[3]National University of „Kyiv-Mohyla Academy“, Kyiv, Ukraine, E-mail: v.lyashko@ukr.net

КВАДРАТИЧНА ЗАДАЧА ДЛЯ МАКСИМАЛЬНОГО k-ПЛЕКСА В НЕОРІЄНТОВАНОМУ ГРАФІ

П. І. Стецюк[1], Т. О. Бардадим[2], В. І. Ляшко[3]

[1]Інститут кібернетики ім. В. М. Глушкова НАН України, Київ, Україна, E-mail: stetsyukp@gmail.com
[2]Інститут кібернетики ім. В. М. Глушкова НАН України, Київ, Україна, E-mail: tbardadym@gmail.com
[3]Національний університет „Києво-Могилянська академія“, Київ, Україна, E-mail: v.lyashko@ukr.net

Abstract. The quadratic optimization problem of finding maximum k-plex in undirected graph is formulated. It is demonstrated that the quadratic problem can be obtained from well-known linear Boolean problem for maximum k-plex. Two families of functionally superfluous quadratic constraints obtained from Boolean problem constraints are reported.
Keywords: maximum k-plex, quadratic optimization problem, Boolean linear programming problem, superfluous constraint.

Резюме. У статті сформульовано квадратичну оптимізаційну задачу для знаходження максимального k-плекса у неорієнтованому графі. Показано, що квадратичну задачу можна отримати з відомої лінійної булевої задачі для максимального k-плекса. Наведено два сімейства функціонально надлишкових квадратичних обмежень, які отримано за допомогою обмежень булевої задачі.
Ключові слова: максимальний k-плекс, квадратична оптимізаційна задача, задача булевого лінійного програмування, функціонально надлишкові обмеження.

Вступ

Структурні властивості графа суттєво залежать від того, чи мають окремі підмножини з множини вершин графа певні, наперед задані властивості. Задачі знаходження максимальних за потужністю таких підмножин мають різноманітні застосування [1]. Окремим випадком таких підмножин є k-плекси. Це поняття для неорієнтованого графа було введене в [2], а задача знаходження максимального k-плекса неорієнтованого графа [2, 3] виникає та активно використовується при аналізі соціальних, телекомунікаційних та інших мереж [4]. При $k = 1$ k-плекс збігається з клікою (повним підграфом) графа. При $k > 1$ k-плекс є ослабленням поняття кліки графа і відповідає слабкішим вимогам на включення вершини в k-плекс, ніж вимоги на включення вершини до кліки.

У статті [3] задачу знаходження максимального k-плекса сформульовано у формі задачі булевого лінійного програмування. Нижче цю задачу сформульовано як квадратичну оптимізаційну задачу та проведено її аналіз, орієнтований на застосування техніки лагранжевих двоїстих оцінок [5].

Послідовність викладу матеріалу буде такою: у розділі 1 наведено загальні відомості про k-плекс, у розділі 2 описано множину допустимих розв'язків для k-плекса графа G за допомогою системи квадратичних обмежень. У розділі 3 розглянуто квадратичну задачу знаходження максимального k-плекса та проаналізовано її зв'язок із булевою лінійною постановкою з [3]. Там же розглянуто сімейства функціонально надлишкових обмежень для уточнення лагранжевих двоїстих оцінок у квадратичній задачі, що базуються на використанні обмежень лінійної булевої задачі для максимального k-плекса.

1. Загальні відомості про k-плекс

Нехай $G=(V,E)$ — неорієнтований граф із множиною вершин $V=\{1,...,n\}$ та множиною ребер E. Ребро графа G, що зв'язує вершини $i \in V$ та $j \in V$, будемо позначати $(i,j) \in E$. Для графа G буде використовуватися також інша форма його представлення: $G = (V, \Gamma)$, де $\Gamma = \{\Gamma(i),\ \ i = 1, ..., n\}$, а $\Gamma(i)$ — кінцеві вершини тих дуг, у яких початковою вершиною є вершина i. Кількість ребер графа G в обох представленнях зв'язані співвідношенням: $|E| = \frac{1}{2}\sum_{i\in V}|\Gamma(i)|$. Комплементарний до G граф будемо позначати $\overline{G} = (V, \overline{E})$ або $\overline{G} = (V, \overline{\Gamma})$, де $(i,j) \in \overline{E}$ і $\overline{\Gamma} = \{\overline{\Gamma}(i),\ \ i = 1, ..., n\}$.

Означення 1. Підмножина вершин S із V називається k-плексом графа G, якщо ступінь кожної вершини в індукованому підграфі $G[S]$ (підграфі, породженому підмножиною S) є не меншою, ніж $|S| - k$.

Підмножина $S \subset V$ є k-плексом, якщо виконується така умова

$$\deg_{G[S]}(i) = |\Gamma(i) \bigcap S| \geq |S| - k \qquad \forall i \in S.$$

k-Плекс є максимальним за включенням (maximal), якщо він не міститься ні в якому іншому k-плексі. Найбільший з максимальних за включенням k-плексів називається максимальним (maximum), його розмір називається k-плексним числом графа G та позначається $\rho_k(G)$ [3]. Очевидно, що 1-плекс є клікою графа G, тому що ступінь кожної вершини в індукованому підграфі $G[S]$ не менше, ніж $|S| - 1$, а це означає, що кожна з вершин у підграфі $G[S]$ зв'язана з усіма іншими вершинами, тобто підграф $G[S]$ є повним підграфом (клікою) графа G. У даному випадку $\rho_1(G) = \omega(G)$, де $\omega(G)$ — клікове число графа G (розмір його максимальної кліки).

Поняття *co*-k-плекса графа G, також введене в [3], є узагальненням поняття незалежної множини вершин графа G. При $k = 1$ *co*-k-плекс збігається з незалежною множиною вершин графа. При $k > 1$ *co*-k-плекс є ослабленням поняття незалежної множини вершин графа (відомої також як внутрішньо стійка множина).

Означення 2. Підмножина вершин S з V називається *co*-k-плексом графа G, якщо виконується така умова

$$\deg_{G[S]}(i) = |\Gamma(i) \bigcap S| \leq k-1 \qquad \forall i \in S.$$

Отже, $S \subset V$ є *co*-k-плексом, якщо ступінь кожної вершини в індукованому підграфі $G[S]$ є не більшою, ніж $k-1$. Очевидно, що *co*-1-плекс є незалежною множиною вершин графа G, оскільки ступінь кожної вершини в індуційованому підграфі $G[S]$ дорівнює нулю, а це означає, що кожна з вершин у підграфі $G[S]$ не зв'язана з жодною з інших вершин підграфа $G[S]$. Відмітимо, що *co*-k-плекс і k-плекс для графа G знаходяться в такому ж зв'язку як кліка графа G і незалежна множина вершин графа G. Тому підмножина S є *co*-k-плексом графа G тоді й тільки тоді, коли S є k-плексом для комплементарного графа $\overline{G}$.

2. Квадратичні обмеження для k-плекса

Нехай вершині $i \in V$ $(i = 1,\ 2, ...)$ відповідає булева змінна $x_i \in \{0, 1\}$ така, що

$$x_i = \begin{cases} 1, & \text{якщо} \quad i \in S, \\ 0, & \text{якщо} \quad i \in V \backslash S. \end{cases}$$

Булеві змінні x_i, $i = 1, \dots, n$ будуть описуватися за допомогою квадратичних обмежень-рівностей

$$x_i^2 - x_i = 0 \qquad \forall i \in V. \tag{1}$$

Побудуємо такі квадратичні обмеження, щоб підмножина S була k-плексом. Ці обмеження повинні задавати вимоги на те, щоб ступінь кожної вершини $i \in S$ у підграфі $G[S]$ була не меншою за $|S| - k$, тобто, щоб у підграфі $G[S]$ кількість дуг, що виходять із кожної вершини $i \in S$ була не меншою за $|S| - k$.

Нехай вершина i належить підмножині S, тобто $x_i = 1$. Позначимо через $N_e(i)$ ступінь вершини i у підграфі $G[S]$, тобто кількість дуг, що виходять з вершини $i \in S$. Тоді в підграфі $G[S]$ ступені вершин із підмножини S задаються за допомогою сімейства співвідношень

$$N_e(i) = \sum_{j \in \Gamma(i)} x_j \quad \forall i \in S. \tag{2}$$

З рівняння $|S| = \sum_{j \in V} x_j$ та умови, що множина S є k-плексом, одержуємо нерівності

$$N_e(i) \geq |S| - k = \sum_{j \in \Gamma(i)} x_j - k \quad \forall i \in S. \tag{3}$$

Із співвідношень (2) та (3) одержуємо сімейство нерівностей

$$\sum_{j\in\Gamma(i)} x_j \geq \sum_{j\in V} x_j - k \quad \forall i \in S, \tag{4}$$

при виконанні яких всі ті вершини $i \in V$, для яких $x_i = 1$, будуть утворювати k-плекс.

Однак, нерівності типу (4) не будуть виконуватися для тих вершин $i \in V$, для яких $x_i = 0$, тобто для всіх $x_i \in V \backslash S$. Для того, щоб одержати квадратичні нерівності, що будуть справедливими і для змінних $x_i = 0$, досить обидві частини нерівності вигляду (4), що відповідає вершині i, помножити на змінну x_i. З огляду на невід'ємність змінної x_i знак нерівності після множення не зміниться, і в результаті одержимо такі нерівності

$$x_i \left(\sum_{j\in\Gamma(i)} x_j \right) \geq x_i \left(\sum_{j\in V} x_j - k \right) \quad \forall i \in V,$$

які можна переписати у вигляді

$$\sum_{j\in\Gamma(i)} x_i x_j \geq \sum_{j\in V} x_i x_j - k x_i \quad \forall i \in V. \tag{5}$$

Квадратичні нерівності (5) разом із обмеженнями (1) повністю описують умови, за яких вершини i належать k-плексу. Дійсно, нерівності (5) будуть справедливі для тих вершин i, для яких $x_i = 1$, оскільки вони переходять в обмеження (4). Нерівності (5) будуть справедливі також для усіх вершин i, для яких $x_i = 0$, тому що вони переходять у тривіальну нерівність $0 \geq 0$.

Зрозуміло, що за допомогою обмежень у вигляді рівностей (1) та у вигляді нерівностей (5) можна описати допустимі булеві розв'язки, які відповідають k-плексу. При цьому зміст обмежень (5) буде пов'язаний із інтерпретацією ступеня вершини, як цього вимагає поняття k-плекса, — в ньому ступінь вершини $i \in S$ більше або дорівнює $|S| - k$. Дійсно, права частина обмеження (5) для вершини i, що належить k-плексу, вказує кількість ребер, що виходять з i-ї вершини, з урахуванням того, що $x_i x_j = 1$ лише тоді, коли обидві змінні x_i та x_j дорівнюють одиниці.

Нерівність (5) можна спростити. Помітимо, що

$$\sum_{j\in V} x_j = \sum_{j\in\Gamma(i)} x_j + \sum_{j\in\overline{\Gamma}(i)} x_j + x_i.$$

Тоді нерівності (5) можна переписати так

$$\sum_{j\in\Gamma(i)} x_i x_j \geq \left(\sum_{j\in\Gamma(i)} x_i x_j + \sum_{j\in\overline{\Gamma}(i)} x_i x_j + x_i^2 \right) - k x_i, \quad \forall i \in V,$$

звідки

$$\sum_{j\in\overline{\Gamma}(i)} x_i x_j \leq k x_i - x_i^2 \quad \forall i \in V.$$

З урахуванням того, що $x_i = x_i^2$ (див. формулу (1)), останні нерівності можна переписати так

$$\sum_{j\in\overline{\Gamma}(i)} x_i x_j \leq (k-1)x_i \quad \forall i \in V. \tag{6}$$

Нерівності (6) разом із рівностями (1) ми покладемо в основу квадратичної моделі для знаходження максимального k-плекса графа G.

Відмітимо, що квадратичним нерівностям (6) можна надати інший зміст, ніж нерівностям (5). Нерівності (6) пов'язані з комплементарним графом $\overline{G}$ і описують таку підмножину вершин S, що ступінь вершини в індуційованому цією підмножиною підграфі $\overline{G}[S]$ не більша за $(k-1)$. Дійсно, для тих вершин $i \in V$, для яких $x_i = 1$, нерівності (6) рівносильні таким нерівностям

$$\sum_{j\in\overline{\Gamma}(i)} x_j \leq (k-1) \quad \forall i \in S,$$

а для тих вершин i, для яких $x_i = 0$, вони рівносильні тривіальним нерівностям $0 \leq 0$. Тому опис підмножини S за допомогою нерівностей (6) та рівностей (1) логічно інтерпретувати як опис *co*-k-плекса для комплементарного графа $\overline{G}$.

3. Квадратична булева задача для $\rho_k(G)$

Ураховуючи, що обмеження (1) та (6) описують множину допустимих варіантів утворення k-плекса, для знаходження максимального k-плекса графа G оптимізаційну квадратичну задачу можна сформулювати в такій формі

$$\rho_k(G) = \max_x \sum_{i\in V} x_i \tag{7}$$

при обмеженнях

$$\sum_{j\in\overline{\Gamma}(i)} x_i x_j \leq (k-1)x_i \quad \forall i \in V, \tag{8}$$

$$x_i^2 - x_i = 0 \quad \forall i \in V. \tag{9}$$

Зрозуміло, що задачу (7)–(9) можна інтерпретувати як задачу знаходження максимального *co*-k плекса графа $\overline{G}$. Із (7)–(9) легко одержати формулювання квадратичної оптимізаційної задачі для *co*-k-плекса графа $\overline{G}$. Для цього досить у сумі лівої частини обмеження (8) замість підсумовування по $j \in \overline{\Gamma}(i)$ використовувати підсумовування по $j \in \Gamma(i)$. Тобто, якщо обмеження (8) замінити на

$$\sum_{j\in\Gamma(i)} x_i x_j \leq (k-1)x_i \quad \forall i \in V, \tag{10}$$

то ми одержимо формулювання квадратичної оптимізаційної задачі знаходження максимального *co*-k-плекса графа $\overline{G}$.

Формулювання задачі (7)–(9) можна одержати із задачі булевого лінійного програмування, запропонованої в [3]:

$$\rho_k(G) = \max_x \sum_{i \in V} x_i \tag{11}$$

при обмеженнях

$$\sum_{j \in \overline{\Gamma}(i)} x_j \le (k-1)x_i + \overline{d_i}(1-x_i) \quad \forall i \in V, \tag{12}$$

$$x_i \in \{0,\ 1\}, \quad \forall i \in V, \tag{13}$$

де $\overline{d_i} = |\overline{\Gamma}(i)|$. Лінійні обмеження (12) побудовано за схемою, подібною тій, що використовувалася у попередньому розділі при переході від обмежень (4), справедливих для $i \in S$ (при $x_i = 1$), до обмежень (5), що є справедливими також для $i \in V \backslash S$ (при $x_i = 0$). Однак, правило для того, щоб обмеження (12) виконувалися для будь-яких $i \in V \backslash S$, тут буде іншим. При $x_i = 1$ обмеження (12) виконуються як нерівності

$$\sum_{j \in \overline{\Gamma}(i)} x_j \le (k-1) \quad \forall i \in S,$$

які повинні бути справедливими для *co*-k-плекса графа $\overline{G}$, що збігається з k-плексом графа G. При $x_i = 0$ обмеження (12) переходять у нерівності

$$\sum_{j \in \overline{\Gamma}(i)} x_j \le \overline{d_i} = |\overline{\Gamma}(i)| \quad \forall i \in V \backslash S,$$

які є справедливими, бо в якості верхньої границі на ступені вершин, що не входять у *co*-k-плекс графа $\overline{G}$, використовується максимально можлива кількість ребер, що виходять з кожної з вершин графа $\overline{G}$. Ті з цих обмежень, де не всі змінні під знаком суми дорівнюють одиниці, будуть надлишковими.

Із задачі лінійного булевого програмування (11)–(13) легко одержати квадратичну задачу (7)–(9). Для цього слід обмеження (13) замінити на відповідний нелінійний аналог (9), а обмеження, яке відноситься до i-ї вершини з (12), помножити на змінну x_i. В силу невід'ємності змінних x_i, $i = 1,\ 2,\ \ldots$ знаки нерівностей при множенні не зміняться, і в результаті одержимо

$$\sum_{j \in \overline{\Gamma}(i)} x_i x_j \le (k-1)x_i^2 + \overline{d_i}(1-x_i)x_i \quad \forall i \in V,$$

звідки, з урахуванням того, що $(1-x_i)x_i = x_i - x_i^2 = 0$ для всіх $i \in V$, приходимо до обмежень (8).

Зрозуміло, що кожна з задач (7)–(9) та (11)–(13) має свої переваги та свої недоліки. Так, наприклад, найсуттєвіша перевага квадратичної задачі над лінійною булевою полягає в тому, що незначним удосконаленням задачі (7)–(9) можна сформулювати квадратичні оптимізаційні задачі знаходження максимальних за розміром підмножин графа із сильнішими властивостями, ніж k-плекс або *co*-k-плекс, — достатньо лише посилити вимогу на включення вершин у ці підмножини. Так, наприклад, умовимося під „строгим“ k-плексом графа G розуміти підмножину його вершин, для яких ступінь вершини дорівнює $|S|-k$. Аналогічно введемо поняття „строгого“ *co*-k-плекса: в ньому ступінь вершини дорівнює рівно $k-1$. Для того щоб сформулювати квадратичні задачі знаходження максимальних із цих підмножин, досить в обмеженнях (8) та (10) замість нерівностей використовувати рівності. Для знаходження „строгого“ k-плекса графа G обмеження (8) слід замінити на обмеження

$$\sum_{j\in\overline{\Gamma}(i)} x_i x_j = (k-1)x_i, \quad \forall i \in V,$$

а для знаходження „строгого“ *co*-k-плекса графа G обмеження (10) слід замінити на такі:

$$\sum_{j\in\Gamma(i)} x_i x_j = (k-1)x_i, \quad \forall i \in V.$$

До переваги лінійної булевої задачі над квадратичною можна віднести той факт, що для задачі (11)–(13) легко підраховувати верхні оцінки для $\rho_k(G)$ за допомогою релаксації обмеження (13). У ряді випадків, наприклад, коли $\rho_k(G)$ буде більше $n/3$, ці оцінки, як правило, можуть виявитися більш ефективними оцінками зверху для $\rho_k(G)$. У результаті для деяких спеціальних графів на базі методу гілок та границь можна реалізувати швидкі алгоритми для знаходження $\rho_k(G)$.

Знаходження верхніх оцінок для квадратичної задачі (7)–(9) є більш трудомістким, ніж для релаксованої задачі (11)–(13). Так, наприклад, якщо в такій якості використовувати лагранжеві двоїсті оцінки [4], [5], то знаходження таких оцінок за допомогою методів недиференційованої оптимізації вимагатиме більше часу, чим у випадку задач лінійного програмування. Однак, для ряду графів лагранжеві двоїсті оцінки можуть виявитися значно точнішими верхніми оцінками, ніж лінійні оцінки. Більш того, існує резерв для уточнення лагранжевих двоїстих оцінок задачі (7)–(9): це введення функціонально надлишкових обмежень [5].

Лінійні обмеження (12) можна використовувати для побудови функціонально надлишкових обмежень з метою покращити точність лагранжевих двоїстих оцінок у багатоекстремальних квадратичних задачах (7)–(9). Якщо скористатися схемою, що використовувалася Н. З. Шором для задачі про максимальну незалежну множину вершин графа [5, с. 250], то при цьому до задачі (7)–(9) додаються два види функціонально надлишкових обмежень. Обмеження першого виду одержано домноженням кожного з лінійних обмежень у (12) на ті змінні x_l, які не входять у це обмеження, тобто додаються $n(n-1)$ надлишкових обмежень вигляду

$$\sum_{j\in\overline{\Gamma}(i)} x_l x_j \le (k-1)x_l x_i + \overline{d_i}(1-x_i)x_l \quad \forall\, i,\, l \in V, \quad i \ne l. \tag{14}$$

Функціонально надлишкові обмеження другого типу можна одержати з лінійних обмежень (12) домноженням на $1-x_l$, $l = 1, 2, ...$. Тут уже можна використовувати $i = l$, бо вони дають нові квадратичні обмеження у формі нерівностей. У результаті маємо n^2 обмежень

$$\sum_{j \in \overline{\Gamma}(i)} x_l x_j - \sum_{j \in \overline{\Gamma}(i)} x_j \leq (k-1) x_l x_i + \overline{d_i}(1 - x_i) x_l \quad \forall\, i,\, l \in V. \qquad (15)$$

Зауважимо, що використання надлишкових обмежень (14), (15) значно збільшує розміри квадратичної задачі. Однак, якщо до попередньої задачі додавати тільки невелику кількість тих надлишкових обмежень, які уточнюють двоїсту оцінку для наступної задачі, то тоді загальна кількість обмежень у кінцевій квадратичній задачі буде невеликою. За такою ж схемою можна використовувати і інші види функціонально надлишкових обмежень для булевих задач, які розглядалися у роботах [6], [7].

Висновки

У роботі побудовано квадратичне формулювання оптимізаційної задачі знаходження максимального k-плекса для неорієнтованого графа. Спорідненість понять k-плекса та кліки дає підстави рекомендувати для її розв'язання підхід, пов'язаний з лагранжевими оцінками та використанням надлишкових обмежень. Цей підхід був запропонований Н. З. Шором та дав ряд важливих теоретичних результатів для задачі про максимальну незалежну множину вершин графа [5].

Роботу виконано за підтримки НАН України, проект 0117U000327.

Література

1. Кристофидес Н. Теория графов. Алгоритмический подход. — Мир: Москва, 1978. — 432 с.
2. Seidman S. B., Foster B. L. A graph theoretic generalization of the clique concept // J. of Math. Sociology. — 1978. — Vol. 6. — P. 139–154.
3. Balansundaram B., Butenko S., Hicks I. V. Clique Relaxations in Social Network Analysis: The Maximum k-plex Problem // Operations Research. — 2011. — Vol. 59, Number 1. — P. 133–142.
4. Нові мережево-орієнтовані методики для інформаційного аналізу великих масивів даних / П. І. Стецюк, М. Г. Журбенко, І. В. Сергієнко та інші // Звіт про науково-дослідну работу М/163-2006, № держ. реєстрації 0106U010005. — Київ: Ін-т кібернетики ім. В.М. Глушкова НАН України, 2006. — 112 с.
5. Shor N. Z. Nondifferentiable Optimization and Polynomial Problems. — London/Boston/Dordrecht: Kluwer Academic Publishers, 1998. — 394 p.
6. Стецюк П. И. О функционально избыточных ограничениях для булевых оптимизационных задач квадратичного типа // Кибернетика и системный анализ. – 2005. – № 6. — С. 168–172.
7. Стецюк П. И. Новые модели квадратичного типа для задачи о максимальном взвешенном разрезе графа // Кибернетика и системный анализ. — 2006. — № 1. — С. 63–75.

Надійшла 28.04.2017