1999 年 4 月 16 日（金） 数理談話会資料

半正定値計画を用いた最大カット問題の .878 近似解法

松井知己
東京大学大学院工学系研究科計数工学専攻

1 はじめに

本稿では，Goemans と Williamson によって近年提案された，組合せ最適化問題の近似解法の解説を行なう．取り上げる話題は，最大カット問題に対する 0.878 近似解法である．

ある解法によって求められる解が，最適解からある程度の距離しか離れていない事が保証できるとき，その解法を近似解法と呼ぶ事が多い．他方，そのような保証ができないとき，その解法を発見的解法と呼ぶ事が多い．しかしながら，上記のような「近似解法」と「発見的解法」の使い分けは，研究者の間でも必ずしも同意されている訳ではない．

目的関数を最大化する最適化問題に対するある解法が，最適値の α 倍程度の目的関数値を持つ許容解を必ず見つけるとき，この解法は α 近似解法であると呼ばれる．ただし α は $0 \leq \alpha \leq 1$ を満たす数である．この定義から明らかなように，最適値が負の値を取る可能性のある問題に対しては，この定義はあまり意味を持たない．

2 最大カット問題

最大カット問題とは以下のような問題である．無向グラフ $G = (V, E)$ と各枝 $e = \{i, j\} \in E$ に対し非負の枝重み $w_e = w_{ij} = w_{ji}$ が与えられているとする．頂点の部分集合 $U \subseteq V$ に対するカット $\delta(U)$ を $\delta(U) = \{\{i, j\} \mid i \in U \not\ni j$ または $i \notin U \ni j\}$ と定義し，カットの重みを $w(U) = \sum_{\{i,j\}\in\delta(U)} w_{ij}$ と定義する．最大カット問題は，重み最大のカットを求める問題である．

最大カット問題は NP- 困難であることが知られている．最大カット問題の 1/2 近似解法は簡単に作ることができる．各頂点を 1/2 の確率で選んだ頂点集合を U とすると，任意の枝に対しその両端点の内ただ 1 つのみが U に入る確率は 1/2 となる．ゆえに，任意の枝に対しその枝が U に対するカットに含まれる確率は 1/2 であり，カットの重みの期待値は枝重みの総和の 1/2 となる．

3 最大カット問題の 0.878 近似解法

本節では d を任意の正整数として議論を進める．以下ではグラフの頂点集合を $V = \{1, 2, \ldots, n\}$ とする．座標空間 R^d の原点を中心とする半径 $1/\pi$ の球の表面を S とする．球面 S 上の任意の 2 点 $\boldsymbol{v}, \boldsymbol{v}'$ に対し，2 点を通る大円の弧のうち短い方の長さを $d(\boldsymbol{v}\boldsymbol{v}')$ とする．

与えられたグラフ G の頂点を球面 S 上に配置する．このときグラフの各枝の重さと距離の積の総和を最大化するような 頂点の配置を求める問題を考える．正確には，以下のように記述される問題である．

$$\begin{array}{lll} \text{MC:} & \text{maximize} & \displaystyle\sum_{\{i,j\}\in E} w_{ij}\ d(\boldsymbol{v}_i\boldsymbol{v}_j) \\ & \text{subject to} & \boldsymbol{v}_i \in S \ \ (\forall i \in V). \end{array}$$

実はこの問題は最大カット問題と本質的に等しい．正確には以下が成り立つ．

定理 1 最大カット問題の最適解（の１つ）を $\delta(U)$ とする．球面 S 上の任意の点 $\boldsymbol{v}^*$ に対し，

$$\boldsymbol{v}_i^* = \begin{cases} \boldsymbol{v}^* & (i \in U), \\ -\boldsymbol{v}^* & (i \in V \setminus U), \end{cases}$$

と定義される解 $(\boldsymbol{v}_1^*, \ldots, \boldsymbol{v}_n^*)$ は問題 MC の最適解である．

証明：頂点配置 $(\boldsymbol{v}_1^*, \ldots, \boldsymbol{v}_n^*)$ に対する問題 MC の目的関数値は，球面 S 上の大円の半円弧の長さが 1 であることから，最大カット問題の最適値に等しい．

以下では， MC の任意の許容解 $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ に対し，その目的関数値が最大カット問題の最適値以下であることを示す．原点を境界に含む閉半空間 H に対し，点集合 $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ のうち H に含まれるものに対応する頂点集合を $V(H)$ とする． また $V(H)$ に対応するカットを $C(H)$, カット $C(H)$ の重みを H に対応するカットの重みと呼ぶ．原点を境界に含む閉半空間 H をすべて等確率で発生させるとき，対応するカットの重みの期待値 z について議論しよう．定義から明らかに， z は最大カットの重み以下である．次に，許容解 $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ の目的関数値が， z となる事を証明する．グラフ中の枝 $\{i, j\}$ がカット $C(H)$ に含まれる確率は， $\boldsymbol{v}_i$ と $\boldsymbol{v}_j$ を結ぶ円弧の 2 倍の長さが，大円に対して占める割合であり， $d(\boldsymbol{v}_i \boldsymbol{v}_j)$ である．ゆえに $z = \sum_{\{i,j\} \in E} w_{ij}\ d(\boldsymbol{v}_i \boldsymbol{v}_j)$ となる． ■

以下では，頂点配置 $(\boldsymbol{v}_1^*, \ldots, \boldsymbol{v}_n^*)$ を最大カット配置と呼ぶ．

問題 MC は最大カット問題と本質的に同じであり, 依然として効率良く解く事は困難である．そこで以下では，2 点 $\boldsymbol{v}_i, \boldsymbol{v}_j$ 間の距離を直線距離（ユークリッド距離）の 2 乗で置き換えた問題について議論する．球面 S 上の任意の 2 点 $\boldsymbol{v}, \boldsymbol{v}'$ に対し， 2 点のユークリッド距離の 2 乗の $\pi^2/4$ 倍を $f(\boldsymbol{v}\boldsymbol{v}')$ と書く，すなわち $f(\boldsymbol{v}\boldsymbol{v}') = (\pi^2/4)||\boldsymbol{v} - \boldsymbol{v}'||^2$ である．ちなみに $f(\boldsymbol{v}\boldsymbol{v}') = (1/2)(1 - \pi^2 \boldsymbol{v}^t \boldsymbol{v}')$ と書くこともできる．この距離は， $d(\boldsymbol{v}\boldsymbol{v}') = 1$ の時， すなわち $\boldsymbol{v}, \boldsymbol{v}'$ が直径の両端となっているとき， $f(\boldsymbol{v}\boldsymbol{v}') = 1$ となるように設定されている．ゆえに，最大カット配置 $(\boldsymbol{v}_1^*, \ldots, \boldsymbol{v}_n^*)$ については， $\displaystyle\sum_{\{i,j\} \in E} w_{ij}\ d(\boldsymbol{v}_i^* \boldsymbol{v}_j^*) = \sum_{\{i,j\} \in E} w_{ij}\ f(\boldsymbol{v}_i^* \boldsymbol{v}_j^*)$ が成り立つ．

ここで以下の問題について考える．

$$\begin{array}{lll} \overline{\mathrm{MC}}: & \text{maximize} & \displaystyle\sum_{\{i,j\} \in E} w_{ij}\ f(\boldsymbol{v}_i \boldsymbol{v}_j) \\ & \text{subject to} & \boldsymbol{v}_i \in S \ \ (\forall i \in V). \end{array}$$

距離 $f(\boldsymbol{v}_i \boldsymbol{v}_j)$ の定義より， $\overline{\mathrm{MC}}$ の最適値は MC の最適値以上である事が分かる．距離 $f(\boldsymbol{v}_i \boldsymbol{v}_j)$ の定義より， 以下の性質が成り立つ事を示す事ができる．

補題 2 球面 S 上の任意の *2* 点 $\boldsymbol{v}, \boldsymbol{v}'$ に対し， $d(\boldsymbol{v}\boldsymbol{v}') \geq \left(\min_{0 \leq \theta \leq \pi} \frac{2}{\pi} \frac{\theta}{1 - \cos\theta} \right) f(\boldsymbol{v}\boldsymbol{v}')$ が成り立つ．またこのとき， $\displaystyle\min_{0 \leq \theta \leq \pi} \frac{2}{\pi} \frac{\theta}{1 - \cos\theta} > 0.878$ である．

証明：ベクトル $\boldsymbol{v}$ と $\boldsymbol{v}'$ のなす角を θ' とすると， $d(\boldsymbol{v}\boldsymbol{v}') = \theta'/\pi$ と $f(\boldsymbol{v}\boldsymbol{v}') = (1 - \cos\theta')/2$ より明らか． ■

上記補題より次の定理が得られる．

定理 3 問題 MC の最適解 $(\boldsymbol{v}_1^*, \ldots, \boldsymbol{v}_n^*)$ および問題 $\overline{MC}$ の最適解 $(\boldsymbol{v}'_1, \ldots, \boldsymbol{v}'_n)$ に対し，

$$\sum_{\{i,j\} \in E} w_{ij}\ d(\boldsymbol{v}'_i \boldsymbol{v}'_j) \geq 0.878 \sum_{\{i,j\} \in E} w_{ij}\ f(\boldsymbol{v}'_i \boldsymbol{v}'_j) \geq 0.878 \sum_{\{i,j\} \in E} w_{ij}\ f(\boldsymbol{v}_i^* \boldsymbol{v}_j^*) = 0.878 \sum_{\{i,j\} \in E} w_{ij}\ d(\boldsymbol{v}_i^* \boldsymbol{v}_j^*)$$

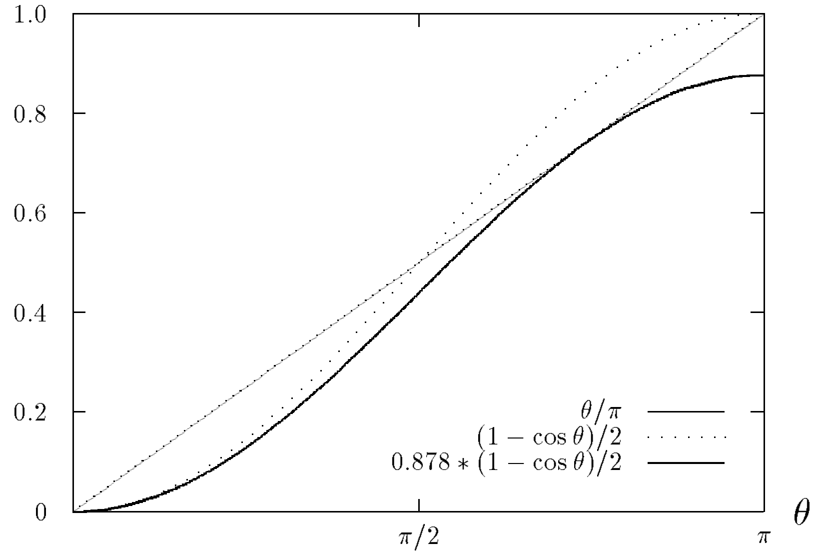

図 1: 関数 θ/π と $(1-\cos\theta)/2$ の関係.

が成り立つ.

上記の定理より，問題 $\overline{\text{MC}}$ を解いて得られる配置では， MC の目的関数値が問題 MC の最適値の 0.878 倍以上となっている．問題 $\overline{\text{MC}}$ を解いて得られる配置に対して，定理 1の証明で用いた手法に従い，原点を境界に含む半閉空間を等確率で発生させて得られるカットの重みの期待値は最大カットの 0.878 倍以上となっている.

3.1 半正定値計画

前節の近似解法を実行するには，問題 $\overline{\text{MC}}$ を解く必要がある．距離 $f(\boldsymbol{v}\boldsymbol{v}')$ の定義から問題 $\overline{\text{MC}}$ は

$$\begin{array}{lll} \overline{\text{MC}}\text{-1:} & \text{maximize} & (1/2)\displaystyle\sum_{\{i,j\}\in E} w_{ij}(1-\pi^2\boldsymbol{v}_i^t\boldsymbol{v}_j) \\ & \text{subject to} & \boldsymbol{v}_i \in S \ \ (\forall i \in V), \end{array}$$

と書くことが出来る，上記の問題は制約式に凸 2 次等式の入った 2 次計画である.

以下では，この問題が半正定値計画に変形される事を示す．変数 y_{ij} $(i,j \in \{1,\ldots,n\})$ を新たに導入し， $y_{ij} = \pi^2\boldsymbol{v}_i^t\boldsymbol{v}_j$ と変数変換を行なうと，行列 $Y = (y_{ij})$ は次の性質；(1) $\forall i \in V,\ y_{ii} = 1$， (2) Y は半正定値対称行列, を満たす．これより以下の問題

$$\begin{array}{lll} \overline{\text{MC}}\text{-2:} & \text{maximize} & (1/2)\displaystyle\sum_{i,j\in E} w_{ij}(1-y_{ij}) \\ & \text{subject to} & y_{ii} = 1 \ \ (\forall i \in V), \\ & & Y = (y_{ij})\text{は半正定値対称行列}, \end{array}$$

は $\overline{\text{MC}}$-1 の緩和問題となる．すなわち $\overline{\text{MC}}$-2 の最適値は $\overline{\text{MC}}$-1 の最適値以上である． $\overline{\text{MC}}$-2 の最適解 Y^* としたとき， $(1/\pi^2)Y^*$ を Cholesky 分解することにより， $(1/\pi^2)Y = X^tX$ が得られたとする．問題 MC を定義し

た座標空間の次元 d を，行列 X の行数とすると，X の列ベクトル集合は問題 $\overline{\mathrm{MC}}$-1 の許容解となり，目的関数は $\overline{\mathrm{MC}}$-2 の最適値と一致する．ゆえに，X の列ベクトル集合は問題 $\overline{\mathrm{MC}}$-1 の最適解である．

上記の半正定値計画には，多項式時間算法があることが知られている．

4 計算機実験

Goemans and Willimason は、計算機実験についても報告をしている. 4 種類のグラフを生成し、計算時間と近似比率を計ったものである. 近似比率は, 実際に求められたカットの重さを, 半正定値計画緩和を用いて得られた上界値で割った値である. カットの生成は, ランダムに 50 個生成して最も良いものを選択している. 半正定値計画問題を解くソフトは, Vandelbei によって作られたものを用いたとある. 実験で確認された近似比率は、すべて 0.9 より大きく, 0.878 よりも大きな値となっている.

表 2と表 3は, 京都大学大学院工学研究科建築学専攻の藤沢克樹氏による実験結果である. 表 2は [3] に記載されたものである. 表 2,3は各サイズのグラフ 1 つについて解いた結果である.「SDP 実行時間」と「メモリ」は, 半正定値計画問題を解くのに要した時間とメモリ量である．RANDOM, SDP-R, LOCAL, TABU という 4 つの解法を実行して, その性能を比較している．RANDOM は, 2 節で紹介した 1/2 近似解法である．SDP-R は 3 節で紹介した 0.878 近似解法である．LOCAL と TABU は、典型的な発見的解法のアイデアを用いた 1/2 近似解法である. 表には, 各々の解法を用いて見つかったカットの重さを, 半正定値緩和問題を解いて得られた上界で割った値を記した．TABU が非常に良い解を得ていることが分かる．RANDOM,LOCAL,TABU の 3 つは、100 秒間実行して得られたカットの中で, 最も良いものを出力として選んでいる．SDP-R は、半正定値計画を解いた後, ランダムなカット生成を 100 秒間行って得られたカットのうち, 最も良いものを出力として選んでいる．SDP-R は、ある程度良い解を出力しているが, 半正定値計画緩和問題を解くのに非常に長い時間がかかっている.

表 3は藤沢克樹氏の御厚意により, 最新の実験結果をお知らせいただいたものを記載している. 数年間で、半正定値緩和問題の求解時間が非常に短くなっていることが分かる.

5 MAX 2SAT 問題の 0.878 近似解法

連言標準形の論理式が与えられているとき，論理式が真となるような基本論理式への真偽の割当があるかを問う問題は，SAT 問題と呼ばれ，NP- 完全であることが証明された最初の問題である．これに対し，与えられた連言標準形の論理式のクローズのうち真となる個数を最大にするような，基本論理式への真偽の割当を求める問題を MAX SAT と呼ぶ．一般形として，各クローズに非負の重みが与えられているとし，真となるクローズの重みの総和を最大化する問題も，MAX SAT と呼ばれる．本稿では，非負重みが与えられた一般形の問題について扱う．

MAX SAT 問題は，各クローズに入っているリテラルの数が k 個以内のとき MAX kSAT 問題と呼ばれる．MAX kSAT 問題は $k=2$ であっても NP- 困難であることが知られている．また MAX 2SAT 問題は，1 未満の定数 α^* が存在して，多項式時間の α^* 近似解法が存在する可能性は，絶望的であることが知られている．

前節の手法は MAX2SAT 問題にも適用できる．まず最初に，制約付き最大カット問題について議論する．無向グラフ $G=(V,E)$ と非負の枝重み w および，枝の部分集合 E' が与えられたとき，E' を含むカットの中で重み最大のものを求める問題を制約付き最大カット問題と呼ぶ．この問題は前節の問題 MC に以下の等式

$$d(\boldsymbol{v}_i\boldsymbol{v}_j) = 1 \ \ (\forall\{i,j\}\in E'),$$

を加えて表される．すると前節と同様の議論を行なうことにより，0.878 近似解法を作ることができる．また最終的に作られる半正定値問題では,

$$y_{ij} = y_{ji} = 1 \ \ (\forall\{i,j\} \in E'),$$

という制約式が付け加わる事となるが，これも効率的に解くことができる．

最後に MAX 2SAT 問題を，上記の制約付き最大カット問題に変形する．基本論理式を $a_1, a_2, \ldots, a_n$ とし，与えられたクローズの集合を $\Gamma = \{C_1, \ldots, C_m\}$ とし，クローズ C_j に与えられた非負の重みを w_j とする．基本論理式 a_i というリテラルには、番号 i を対応させ、基本論理式 a_i の否定というリテラルには、番号 $n+i$ を対応させる。各クローズは，それに含まれるリテラルに対応する番号の集合と見なす.

このとき完全グラフ $G = (V, E)$ と枝集合 E' を

$$V = \{0, 1, 2, \ldots, n, n+1, \ldots, 2n\}, \ \ E' = \{\{i, n+i\} \mid i = 1, 2, \ldots, n\},$$

と定義し，さらに枝 $\{i,j\} \in E$ について $i < j$ のとき枝重み w_{ij} を

$$w_{ij} = \begin{cases} (1/2) \displaystyle\sum_{C_l \supseteq \{i,j\}} w'_l & (\text{if } 0 < i), \\ (1/2) \displaystyle\sum_{C_l \ni j} (3 - |C_l|) w'_l & (\text{if } i = 0), \end{cases}$$

と定義する．上記の枝重み $w(e)$ は次の手続きでも作ることが出来る．まず最初は全ての枝重みを 0 とする．各クローズ C_l について，(1) $C_l = \{i\}$ すなわちリテラルが 1 個ならば，枝 $\{0, i\}$ の重みを w_l 増やす，(2) $C_l = \{i, j\}$ すなわちリテラルが 2 個ならば，頂点 $\{0, i, j\}$ からなる長さ 3 のサイクル上の枝の重みをそれぞれ $(1/2)w_l$ 増やす.

このとき E' を含む任意のカットを $\delta(S)$ とし，一般性を失うこと無く $0 \in S$ と仮定すると，以下が成り立つ。頂点集合 $V \setminus S$ に含まれる頂点に対応する基本論理式を真としたとき，真となるクローズの重みの合計はカットの重みに等しい．逆に任意の真偽の割当に対して，上記の逆の手続きで得られるカットの重みは，真となるクローズの重みの合計に等しい．ゆえに上記の手続きで MAX 2SAT 問題を制約付き最大カット問題に変形される。

参考文献

[1] M. X. Goemans and D. P. Williamson (1994). “.878-approximation Algorithms for MAX CUT and MAX 2SAT,” Proceedings of 26th STOC, Montreal, Quebec, Canada, 422–431.

[2] M. X. Goemans and D. P. Williamson (1995), “Improved Approximation Algorithms for Maximum Cut and Satifiability Problems Using Semidefinite Programming,” J. ACM 42, 1115–1145.

[3] 藤沢克樹 (1996), “組合せ最適化問題に対する近似解法,” 第 8 回 RAMP シンポジウム論文集，東京大学, 139–154.

表 1: 0.878 近似解法の計算実験 : Vanderbei : Sun SPARC Station 1

問題種類	頂点数	問題数	平均近似比率	平均時間 [秒]
A	50	50	0.96988	36.28
	100	20	0.96783	323.08
	200	5	0.97209	4629.62
B	50	50	0.97202	23.06
	100	20	0.97097	217.42
	200	5	0.97237	2989.00
C	50	50	0.95746	23.53
	100	20	0.94214	306.84
	200	5	0.92362	2546.42
D	50	50	0.95855	27.35
	100	20	0.93984	355.32
	200	5	0.93635	10709.42

表 2: 計算機実験 (1996 年 8 月) : SDPA Ver 1.0 : Sun SPARC Station 20 : Super SPARC II 125MHz

頂点数	枝密度 [%]	SDP 実行時間 [秒]	メモリ [Mbyte]	RANDOM	SDP-R	LOCAL	TABU
124	2	636.5	5.1	0.6761	0.9648	0.9507	0.9648
124	4	610.3	5.1	0.7040	0.9374	0.9411	0.9485
124	8	637.1	5.1	0.7781	0.9470	0.9513	0.9534
124	16	607.0	5.1	0.8052	0.9625	0.9648	0.9648
250	1	7948.5	23.5	0.6303	0.9518	0.9140	0.9612
250	2	7407.6	23.5	0.6612	0.9516	0.9118	0.9438
250	4	7290.7	23.5	0.7318	0.9244	0.9376	0.9448
250	8	7505.7	23.5	0.7712	0.9454	0.9513	0.9567

表 3: 計算機実験 (1999 年 4 月) : SDPA Ver 4.30 : DEC ALPHA 21164 600MHz

頂点数	枝密度 [%]	SDP 実行時間 [秒]	メモリ [Mbyte]
124	2	9.1	5.0
124	4	9.4	5.0
124	8	9.4	5.0
124	16	9.5	5.0
250	1	82.7	20.0
250	2	79.8	20.0
250	4	80.3	20.0
250	8	80.2	20.0