3.1　離散・連続ハイブリッド最適化

室田 一雄，松井 知己，岩田 覚
情報理工学系研究科 数理情報学専攻

1　M 凸劣モジュラ流問題の解法

室田 一雄（数理情報学専攻）

1.1　概要

M凸関数は, 整数格子点上で定義された関数のクラスとして, 離散凸解析 [7] において中心的な役割を担っている．M凸劣モジュラ流問題は, 効率良く解くことができる組合せ最適化問題の最も一般的な枠組みの一つであり, その特殊ケースとして最小費用流問題や劣モジュラ流問題を含み, 数理経済学への応用もある．本研究では，この問題に対する効率的なアルゴリズムを構成した．とくに，M 凸関数がスケーリングに閉じているという意味で構造的にロバストであれば，弱多項式時間の容量スケーリングアルゴリズムが構成できることを示した [5].

1.2　M 凸関数

M凸関数の定義を述べる．V を有限集合とする．関数 $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ が**M凸関数**であるとは, f が**交換公理** (M-EXC) を満たすことである．

(M-EXC) $\forall x, y \in \mathrm{dom}\, f$, $\forall u \in \mathrm{supp}^+(x-y)$, $\exists v \in \mathrm{supp}^-(x-y)$:
$f(x) + f(y) \geq f(x - \chi_u + \chi_v) + f(y + \chi_u - \chi_v)$.

ここで, $\mathrm{dom}\, f = \{x \in \mathbf{Z}^V \mid f(x) < +\infty\}$,

$$\mathrm{supp}^+(x-y) = \{w \in V \mid x(w) > y(w)\},$$
$$\mathrm{supp}^-(x-y) = \{w \in V \mid x(w) < y(w)\}$$

であり, $\chi_w \in \{0,1\}^V$ は $w \in V$ の特性ベクトルとする．

正の整数 α に対して, $f^\alpha : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ を

$$f^\alpha(x) = f(\alpha x) \qquad (x \in \mathbf{Z}^V)$$

と定義する．この操作を**スケーリング**と呼ぶ．一般に f がM凸関数であっても f^α はM凸関数とは限らない．しかし, 分離凸関数や 2 次のM凸関数, 層凸関数などのクラスのM凸関数はスケーリングについて閉じている．

1.3　M凸劣モジュラ流問題

M凸劣モジュラ流問題とは以下のように定義される問題である．有向グラフ $G = (V, A)$ を考える．各枝の流量の下限と上限 $\underline{c}, \overline{c} \in \mathbf{Z}^A$ および, 単位流量あたりの費用 $\gamma \in \mathbf{R}^A$ が与えられているとする．容量の上下限制約を満たし与えられた供給量と整合的なフロー $\xi \in \mathbf{Z}^A$ の中で総費用を最小にするものを求める問題である最小費用流問題は, 供給制約をフローの境界 $\partial\xi$ がある基多面体に属すべきであるという制約に置き換えた劣モジュラ流問題に一般化される．劣モジュラ流問題はポテンシャル (双対変数) よる最適性規準, 負閉路による最適性規準, 最適解の整数性, 効率的なアルゴリズムの存在などの良い性質をもっており, 扱いやすい組合せ最適化問題の代表となっている．

さらに, $\partial\xi$ に対するコスト関数 $f : \mathbf{Z}^V \to \mathbf{R} \cup \{+\infty\}$ を目的関数に加えるという一般化を考えるとき, f がM凸関数ならば良い性質が保たれる．M凸関数 f によって以下のように定義される問題を整数値フローの**M凸劣モジュラ流問題**と呼ぶ．

$$\begin{array}{ll} \text{minimize} & \Gamma_2(\xi) = \displaystyle\sum_{a \in A} \gamma(a)\xi(a) + f(\partial\xi) \\ \text{subject to} & \underline{c}(a) \leq \xi(a) \leq \overline{c}(a) \quad (a \in A), \\ & \partial\xi \in \mathrm{dom}\, f, \\ & \xi(a) \in \mathbf{Z} \quad (a \in A). \end{array}$$

1.4　アルゴリズム

本研究では，この問題に対するアルゴリズムとして，最短路繰返し法を与え，さらに最短路繰返し法にスケーリング技法を用いて, スケーリングについて閉じたM凸関数 f による MSFP に対する効率的な算法である容量スケーリング法を構成した.

2　多項式時間サンプリング法の提案

松井 知己（数理情報学専攻）

2.1　概要

マルコフ連鎖は，マルコフ連鎖モンテカルロ法や，ベイズ推定に用いられる手法であり，金融，バイオインフォマティクス，統計的検定理論，音声認識，画像復元等様々な分野で用いられている．特に，マルコフ連鎖を用いたサンプリング手法において，サンプルを得るまでの必要推移回数と，要求精度の関係について算定を行うことで，現実問題に適用できる十分な速度を持つ算法の設計が要求される．多くの応用では，サンプリングのための反復回数は経験的な手法で決定されているが，重要な問題では理論的な保証が必要となる．

マルコフ連鎖を用いてサンプリングする際，サンプリングするまでの推移回数の算定は，経験的に定められていることが多い．これに対し，近年，マルコフ連鎖を用いたサンプリング法に対する新たな理論と手法が生れつつあり，実際問題への適用が予想される．収束の速さを理論的に見積る手法の研究には，path coupling, canonical path method, conductance bound 等が開発されている．これらの手法を用いるには，対象とするマルコフ連鎖毎に，その適用方法を提案する必要がある．また，coupling from the past (CFTP) と呼ばれる算法は，有限時間で定常分布からサンプルできる画期的手法であるが，これを用いるには様々な前提条件が満される必要がある．現実に用いられているマルコフ連鎖にこれらの手法を用いる研究が要求されている．

医療統計の分野で検定に用いられる分割表において，これに対し，どんな数値例に対しても必ず要求精度を守る解を生成する高速な近似解法を提案している．医療統計の分野でもちいられる分割表において，正確検定を行う際，マルコフ連鎖モンテカルロ法 (MCMC 法) による積分計算が用いられる事が多い．特に分割表の総数を数え上げる事は重要かつ難しい問題であり，近年様々な手法が開発されつつある．

本研究では，現実にしばしば現れる形状の多次元分割表に対し，MCMC 法に用いる高速なマルコフ連鎖の提案を行った．また，２次元分割表に対しその総数を数える算法について，既往の算法によって得られる値が偏る事を計算実験によって示し，偏りの少ない解が得られる算法の提案を行っている．

2.2　提案する算法

論文 [3] では, 医療統計の分野で検定に用いられる分割表において，その数を数え上げる問題に対し，どんな数値例に対しても必ず要求精度を守る解を生成する高速な近似解法を提案している．この解法の速度は，多項式時間サンプリング算法が存在するという仮定の下では，分割表の入力サイズと，要求精度の逆数の多項式時間である．多項式時間サンプリング算法の存在については不明であり，世界で数多くの研究者によって現在挑戦されている有名な問題である．

論文 [4] では，分割表をサンプルする多項式時間近似算法の提案を行っている．医療統計において正確検定を行う際，マルコフ連鎖モンテカルロ法 (MCMC 法) による積分計算が用いられる事が多い．本論文では，現実にしばしば現れる形状の多次元分割表に対し，MCMC 法に用いる高速なマルコフ連鎖の提案を行った．本研究で提案するサンプリングアルゴリズムは，確率的近似解法であり，その計算時間は，分割表の入力サイズと近似比率の入力サイズの多項式で抑えられる．多次元分割表に対し，多項式時間の近似サンプリング算法は現在までに殆ど知られておらず，この結果から今後さらなる発展が期待される．

2.3 おわりに

バイオインフォマティクスの分野では，配列解析に用いる統計的手法に対し，マルコフ連鎖を用いた手法が多数提案されている．これらの手法に対し，その実行に必要な時間を理論的に保証することが期待されている．現在，配列解析に用いられるベイズ的手法において重要とされる Dirichlet 分布に対し，多項式時間算法の構築にほぼ成功しており，現在その効果を計算実験で確認中である．

今後は，ポストゲノム解析の分野，特にオーダーメイド治療に必要なハプロタイプ解析において，標準的な算法に必ず組込まれるような算法の開発が期待される．更に音声認識や文字認識，自動翻訳等の，隠れマルコフモデルを用いた推定問題に対し，その解法の高速化が期待される．また，画像復元やデリバティブの価格計算等，マルコフ連鎖を用いた手法が成功をおさめている巨大問題に対し，高速かつ高性能の解を与える算法の開発につながる事が期待される．

3 行列束に関するロバスト線形計算

岩田 覚（数理情報学専攻）

微分代数方程式で記述される動的システムの解析においては, 行列束の Kronecker 標準形の構造指数が基本的な役割を果たしている．しかし, 構造指数は, 摂動に対して不安定な離散値を取るため, 正確に計算するには, 計算誤差に対してロバストな手法の開発が必要とされる. 本研究は, 行列束の零非零パターンに代表される離散構造に注目し, 離散最適化の高度な技法を積極的に活用することによって, この様な要請に応えることを目的としている.

行列束とは, 同じサイズの定数行列の組 (A, B) のことであり，各要素が 1 次以下の多項式行列 $M(s) = sA + B$ として表現される． 任意の行列束 $M(s)$ に対して，正則行列 P，Q が存在して，$K(s) = PM(s)Q$ が $\{sI_\nu + J, N_{\mu_1}, \cdots, N_{\mu_b}, L_{\epsilon_1}, \cdots, L_{\epsilon_p}, U_{\eta_1}, \cdots, U_{\eta_q}\}$ を成分とするブロック対角行列となる． ここで，I_ν は ν 次単位行列であり, N_μ は μ 次正則行列束

$$N_\mu = \begin{bmatrix} 1 & s & 0 & \cdots & 0 \\ 0 & 1 & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & s \\ 0 & \cdots & \cdots & 0 & 1 \end{bmatrix},$$

L_ϵ は $\epsilon \times (\epsilon + 1)$ 行列束

$$L_\epsilon = \begin{bmatrix} 1 & s & 0 & \cdots & 0 \\ 0 & 1 & s & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & s \end{bmatrix}$$

を表す. また, U_η は L_η の転置行列である. ブロック対角行列 $K(s)$ を, $M(s)$ の Kronecker 標準形, $\nu, \mu_1, \cdots, \mu_b, \epsilon_1, \cdots, \epsilon_p, \eta_1, \cdots, \eta_q$ を, その構造指数と呼ぶ. 特に, $\mu_1, \cdots, \mu_b$ を巾零指数, $\epsilon_1, \cdots, \epsilon_p$ を列指数, $\eta_1, \cdots, \eta_q$ を行指数と呼ぶ.

巾零指数は, $M(s)$ の k 次小行列式の最大次数 $\delta_k(M)$ を各 k に対して計算することによって, 決定することができる. 一般の多項式行列に対して, $\delta_k(M)$ を計算する組合せ緩和算法が室田 [6] によって提案された. この算法は, 零非零パターンから得られる 2 部グラフ上のマッチング問題を解いて, $\delta_k(M)$ の推定値を得, この推定値が妥当か否かを数値計算を用いて検証し, 妥当でないと判定された場合には行列を変形するという操作を繰り返す.

計算機実験の結果, この組合せ緩和算法は, 計算時間の点では効率的であるものの, 多項式行列の変形を行った結果, 計算に必要とされる記憶容量が大きくなるという問題点を抱えていることが明らかとなった [2]. そこで, 対象を行列束に限定した上で, 定数行列による同値変換のみを用いた新たな組合せ緩和算法を開発し, 必要となる記憶容量を格段に減らした [1]. この結果, 誤差を伴う数値計算を極力避けて, 行列束の巾零指数を正確に計算する効率的なアルゴリズムが得られた.

さらに, 行指数/列指数に関する組合せ緩和法の開発を目標として, 数値計算を全く行わずに, 行指数, 列指数の推定値を得る組合せ的な手法を導

入した．この手法では，巾零指数の場合と異なり，非零要素が代数的独立であると仮定した場合にも，推定値が必ずしも真値と一致しないことが明らかとなった．今後は，この点に留意しつつ，組合せ緩和算法の研究を進めることを計画している．

参考文献

[1] S. Iwata: Computing the maximum degree of minors in matrix pencils via combinatorial relaxation, *Algorithmica*, to appear.

[2] S. Iwata, K. Murota and I. Sakuta: Primal-dual combinatorial relaxation algorithms for the maximum degree of subdeterminants, *SIAM J. Sci. Comput.*, **17** (1996), 993–1012.

[3] S. Kijima and T. Matsui: Approximate counting scheme for $m \times n$ contingency tables, *Technical Report*, METR 2003-01, Dept. Math. Eng. Info. Phys., University of Tokyo, January 2003.

[4] T. Matsui, Y. Matsui, and Y. ONO: Random generation of $2 \times 2 \times \cdots \times J$ contingency tables, *Technical Report*, METR 2003-03, Dept. Math. Eng. Info. Phys., University of Tokyo, January 2003.

[5] S. Moriguchi and K. Murota: Capacity scaling algorithm for scalable M-convex submodular flow problems, *Optimization Methods and Software*, to appear.

[6] K. Murota: Combinatorial relaxation algorithm for the maximum degree of subdeterminants: Computing Smith-McMillan form at infinity and structural indices in Kronecker form, *Appl. Algebra Engin. Comm. Comput.*, **6** (1995), 251–273.

[7] 室田一雄: 離散凸解析, 共立出版, 2001.