離散・連続ハイブリッド最適化

室田一雄 * 松井知己 * 岩田覚 * 松尾宇泰 † 土谷隆 ‡

*情報理工学系研究科数理情報学専攻
†情報学環, 情報理工学系研究科数理情報学専攻
‡統計数理研究所, 情報理工学系研究科数理情報学専攻

1 離散ヘッセ行列と局所 2 次展開

1.1 概要

連続変数の凸解析において, 適当な微分可能性を仮定すると, 関数 g に対して

$$
\begin{aligned}
& g \text{ が凸関数} \\
& \quad \Longleftrightarrow \quad g \text{ のヘッセ行列が各点で半正定値} \\
& \quad \Longleftrightarrow \quad g \text{ の局所 2 次近似が各点で凸}
\end{aligned}
$$

が成り立つ. 本研究 [1] では, L 凸関数 [2, 3] に対して, この同値性の離散版を構築する. すなわち, 離散ヘッセ行列と局所 2 次展開を定義し, これらを用いて L 凸関数の特徴付けを行う.

一般に, 離散関数 g に関する離散ヘッセ行列 H に下記の性質を要請するのは自然であろう.

- H は g の局所的な情報により定義される対称行列.
- g がアフィン関数のとき H は零行列.
- H は g について線形.
- $g(x) = \frac{1}{2}x^\top Ax$ のとき H は A に一致.

離散関数のクラス $\mathcal{C}$ が与えられたとき, 離散ヘッセ行列を適当に定義し, さらに, 以下のような性質を持つ行列のクラス $\mathcal{H}$ を見出すことによって, 離散ヘッセ行列のクラスの特徴付けを試みる.

- 行列が $\mathcal{H}$ に属する $\Longleftrightarrow$ その行列が, $\mathcal{C}$ に属する関数のある点における離散ヘッセ行列である.
- 関数が $\mathcal{C}$ に属する $\Longleftrightarrow$ その関数の離散ヘッセ行列が, 各点で $\mathcal{H}$ に属する.
- 関数が $\mathcal{C}$ に属する $\Longleftrightarrow$ その関数の離散ヘッセ行列による局所 2 次近似が, 各点で $\mathcal{C}$ に属する.

本研究で提案する L 凸関数に対する離散ヘッセ行列はこれらの性質を満たしている.

1.2 結果

離散関数 $g : \mathbf{Z}^n \to \mathbf{R}$ と $p \in \mathbf{Z}^n$ に対して, 離散ヘッセ行列 $H(g;p)$ を以下で定義する:

$$
\begin{aligned}
H_{ij}(g;p) &= -\eta_{ij}(g;p) \quad (i \neq j), \\
H_{ii}(g;p) &= \eta_i(g;p) + \textstyle\sum_{j \neq i} \eta_{ij}(g;p).
\end{aligned}
$$

ただし, χ_i を第 i 単位ベクトルとして,

$$
\begin{aligned}
\eta_{ij}(g;p) &= -g(p + \chi_i + \chi_j) \\
&\quad + g(p + \chi_i) + g(p + \chi_j) - g(p), \\
\eta_i(g;p) &= g(p) + g(p + \mathbf{1} + \chi_i) \\
&\quad - g(p + \mathbf{1}) - g(p + \chi_i).
\end{aligned}
$$

次の定理は L 凸関数に対する離散ヘッセ行列のクラスを特徴付ける.

定理 次の 3 条件は同値である.

(a) g は L 凸関数.

(b) 各 p に対して, $H_{ij}(g;p) \leq 0$ $(\forall i \neq j)$ かつ $\sum_j H_{ij}(g;p) = 0 \quad (\forall i)$.

(c) 各 p に対して, $\hat{g}(q;p)$ は q に関する L 凸関数.

2 パーフェクトサンプリング法

2.1 概要

マルコフ連鎖を用いたサンプリングは，MCMC 法（マルコフ連鎖モンテカルロ法）等における重要な問題である．従来のサンプリング法では，初期状態から十分な回数推移させることによりサンプリングを行っているが，その回数の算定は経験に基づくものであることが多い．本研究で対象とする「マルコフ連鎖を用いたパーフェクトサンプリング法」は，目的とする分布を極限分布に持つマルコフ連鎖から，極限分布に「厳密に」従うサンプルを得る手法である．

従来手法では，(予め定めた) 有限回の推移の後，サンプリングを行うため，その回数をどれほど多くしても「厳密な」サンプリングは不可能である．1996 年に Propp と Wilson によって提案された coupling from the past 法 (以下 CFTP 法) は，マルコフ連鎖の極限分布に厳密に従うサンプリングを可能とする驚くべき手法である [4].

本研究は，ゲノム解析や検定理論，非線形離散最適化等の応用分野に現れる分布に対する CFTP 法の開発を目指しており，これまでの理論的研究とは一線を画すものである．本研究では，離散化 Dirichlet 分布に対する CFTP 法の開発を行った．離散化 Dirichlet 分布は，ゲノム解析等の隠れマルコフモデルで多用される分布であり，高速なサンプリング法が必要とされている．これに対し，CFTP に基づくパーフェクトサンプリング法を構築し，その多項式時間性について示した．

2.2 結果

Dirichlet 分布は，ゲノム解析等の隠れマルコフモデルにおいて，多項分布の事前分布として用いられる．本来の Dirichlet 分布は，高次元の開単体上に定義される連続分布であるが，計算機シミュレーションによってサンプリングを行う場合は，定義域を離散化して得られる離散 Dirichlet 分布が用いられる．離散 Dirichlet 分布に対しては，Matsui, Motoki, Kamatani [5] によって，多項式時間近似サンプリング法が提案されている．

本研究では，離散 Dirichlet 分布を唯一の定常分布に持つ特殊な hit and run chain を構成し，そのマルコフ連鎖の単調性を示すことに成功した．定義域を離散化するメッシュサイズを $1/\Delta$ とし，Dirichlet ランダムベクトルの次元を n とすると，提案する単調 CFTP 法における推移の回数の期待値は $\mathrm{O}(n^3 \ln \Delta)$ となることを示した．

この解法の長所の一つは，計算時間が Dirichlet パラメータの大きさに依存しないことであり，この点において従来の棄却サンプリング法に優っている．また，従来の近似サンプリング法と違い誤差パラメータを自分で設定する必要がない．

3 符号対称行列の Sylvester 指数

3.1 概要

社会現象を線形方程式系で定式化する際には，選好順序や係数の正負などの大小関係は容易に知ることができるが，係数の数値化が困難である場合が多い．このような定性的情報のみ与えられている状況で解析を行うためには，係数行列が要素の絶対値によらず正則 (符号正則) かどうかを知る必要が生じる．この符号正則性を判定する問題の計算複雑度は長い間未解決であったが，1999 年に Robertson, Seymour, Thomas [7] によって多項式時間解法が示された．本研究では, 符号正則な対称行列の Sylvester 符号指数を決定する多項式時間アルゴリズムを提案した．

n 次対称行列 A に対し，適当な正則行列 S による合同変換で

$$S^\top AS = \mathrm{diag}(1, \ldots, 1, -1, \ldots, -1, 0, \ldots, 0)$$

という対角行列に変換したときの $1, -1, 0$ の個数が Sylvester 符号指数である．Sylvester 符号指数は, 2 次形式の概形を規定する基本的な不変量である．

対称行列 A の行列式の展開項で非零なものが存在する時，A は項別正則であると言う．正則な

対称行列は項別正則である．また, 非零要素の絶対値によらず, 同じ符号要素の対称行列が全て正則であるとき, 符号正則であるという．

符号正則な対称行列に対しては，Sylvester 符号指数が行列要素の符号情報のみから一意に定まることが Hall, Li, Wang [6] によって明らかにされた．本研究では, 項別正則な対称行列の組合せ的構造を解析することによって, 符号指数を効率的に計算するアルゴリズムを設計した．

3.2 結果

n 次正方行列 A を表現する 2 部グラフ $G(A) = (U, V; E)$ を，頂点集合 $U := \{u_1, \ldots, u_n\}$, $V := \{v_1, \ldots, v_n\}$, 枝集合 $E := \{(u_i, v_j) \mid a_{ij} \neq 0, u_i \in U, v_j \in V\}$ で定める．各頂点に高々一つの枝が接続している枝集合をマッチング，そして全ての頂点に接続している時を完全マッチングという．正方行列 A に対し，$G(A)$ が完全マッチングを持つ時，かつその時に限り A が項別正則である．

項別正則な対称行列 A に対応する 2 部グラフ $G(A) := G(U, V; E)$ の構造を調べる．項別正則性から $G(A)$ は完全マッチング M を持つ．枝集合 $X \subseteq E$ に対し，X の転置 $X^\top := \{(u_j, v_i) \mid (u_i, v_j) \in X\}$ を定義する．すると対称性より完全マッチング M の転置 $M^\top$ は完全マッチングであり，サーキット $C \subseteq E$ の転置 $C^\top$ はサーキットである．$C^\top$ と C が等しい時, 対称なサーキットと呼ぶ．両端点がサーキット C 上にあり C に含まれない枝を C の弦と呼ぶ．この時次の定理が成り立つ．

定理　項別正則な対称行列 A に対し，$G(A)$ は以下のいずれかの条件を満たす．

- $a_{ii} \neq 0$ となる (u_i, v_i) を含む完全マッチングが存在する．
- $a_{ij} = a_{ji} \neq 0$ $(i \neq j)$ となる (u_i, v_j), (u_j, v_i) を含む完全マッチングが存在する．
- $G(A)$ は弦のない対称なサーキットの直和．

この定理を用いて, 符号正則な n 次対称行列の符号指数を, 行列要素の符号情報のみから計算するアルゴリズムを設計した．その計算量は, 非零要素数を m とするとき, $O(nm)$ である．一方で, 項別正則とは限らない一般の対称行列に関しては, 行列要素の符号情報のみから符号指数が決定不可能かどうかを判定する問題は，NP 完全であることも明らかとなった．

4 高精度な保存・散逸数値解法

4.1 概要

保存系，あるいは散逸系をなす常微分方程式系を数値的に解く際は，その保存・散逸性を離散系でも保つことが望ましい．しかし既存の保存・散逸数値解法は，一部の特殊な場合を除いてすべて時間刻み幅に対して 2 次精度であった．一般的な保存・散逸系に対して，高精度な保存・散逸数値解法を初めて与えたのが Matsuo [8] であるが，これも高々6 次までという制限があった．

本年度の研究では，Matsuo [8] を改良し，理論的には任意の次数まで安定な解法を構成できることを示した．この方法は，Brugnano, Trigiante [9] による常微分方程式に対する境界値法を援用するもので，プログラム実装と計算コストの面で欠点を負う代わりに高い安定性を得る．

4.2 対象とする系と既存の手法

方程式：

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{z}(t) = A\nabla H(\boldsymbol{z}), \quad t > 0, \\ \boldsymbol{z}(0) = \boldsymbol{z}_0 \end{cases} \tag{1}$$

を考える．但し $\boldsymbol{z} : \mathbf{R} \to \mathbf{R}^N$，$A$ は $N \times N$ の実行列，$H : \mathbf{R}^N \to \mathbf{R}$，$\nabla H(z)$ は H の勾配，そして $\boldsymbol{z}_0 \in \mathbf{R}^N$ は初期値である．行列 A が歪対称の時，(1) は保存系となる：

$$\frac{\mathrm{d}}{\mathrm{d}t}H(\boldsymbol{z}(t)) = (\nabla H)^{\mathrm{T}}\dot{\boldsymbol{z}} = (\nabla H)^{\mathrm{T}} A\nabla H = 0.$$

Hamilton 系はこの場合に属す．一方，行列 A が半負定値の場合，(1) は散逸系となる：

$$\frac{\mathrm{d}}{\mathrm{d}t}H(\boldsymbol{z}(t)) = (\nabla H)^{\mathrm{T}} A\nabla H \leq 0.$$

例えば Hamilton 系に散逸項を付けた系がこの場合に属す.

Matsuo [8] は，(1) に対する保存・散逸解法として次を提案した.

$$\delta_{m;\boldsymbol{c}}^{\langle 1\rangle,p}\boldsymbol{z}^{(m)} = A\nabla_{\mathrm{d};\boldsymbol{c}}^{p}H(\boldsymbol{z}^{(m-l_1)},\ldots,\boldsymbol{z}^{(m+l_2)}) \quad (2)$$

ただし $\delta_{m;\boldsymbol{c}}^{\langle 1\rangle,p}$ はある p 次差分作用素を，$\nabla_{\mathrm{d};\boldsymbol{c}}^{p}$ は $\delta_{m;\boldsymbol{c}}^{\langle 1\rangle,p}$ に応じて定まる離散勾配を表す．エネルギー H も，それに応じて多段に拡張されている．作用素 $\delta_{m;\boldsymbol{c}}^{\langle 1\rangle,p}$ は，安定性の要請から通常後退差分作用素にとられる．つまりスキーム (2) はほぼ,（保存・散逸に限らない）一般的な数値解法，後退差分公式 (BDF) と同じ形であり，それと同じ制限（7 次以上は常に不安定）を受ける.

4.3　新しい保存・散逸解法

Brugnano, Trigiante [9] は，BDF の安定性を高めるために一般化 BDF (GBDF) を提案した.要点は次の 3 点である: (1) 局所スキームとして後退差分公式と異なる，より対称性の高いスキームを用いる，(2) そのスキームを，初期値からの逐次ではなく，初期値と最終値を定めた上で大域的に連立して解く（→境界値法），(3) 存在しない最終値の分は，別途方程式を立てて辻褄を合わせる.

本研究ではこの考え方を Matsuo [8] に拡張適用することにより，GBDF の高い安定性を保ったまま，保存・散逸性をもつスキームを構成した.安定性は，実際いくつかの数値実験で確認されている．一方でスキーム全体を大域的に解くコストは軽くなく，そのコストを支払っても 7 次以上の精度が必要とされるかを，利用者が吟味する必要がある.

参考文献

[1] S. Moriguchi and K. Murota: Discrete Hessian Matrix for L-convex Functions, *IEICE Trans. Fundamentals of Electronics, Communications and Computer Sciences*, to appear.

[2] 室田一雄: 離散凸解析, 共立出版, 2001.

[3] K. Murota: *Discrete Convex Analysis*, SIAM, 2003.

[4] J. Propp and D. Wilson: Exact Sampling with Coupled Markov Chains and Applications to Statistical Mechanics, *Random Structures and Algorithms*, **9** (1996), 232–252.

[5] T. Matsui, M. Motoki, and N. Kamatani: Polynomial Time Approximate Sampler for Discretized Dirichlet Distribution, *Lecture Notes in Computer Science*, Springer-Verlag, **2906** (2003), 676–685.

[6] F. J. Hall, Z. Li, and D. Wang: Symmetric Sign Pattern Matrices That Require Unique Inertia, *Linear Algebra Appl.*, **338** (2001), 153–169.

[7] N. Robertson, P. D. Seymour, and R. Thomas: Permanents, Pfaffian Orientations, and Even Directed Circuits, *Ann. Math.*, **150** (1999), 929–975.

[8] T. Matsuo: High-order Schemes for Conservative or Dissipative Systems, *J. Comput. Appl. Phys.* **152** (2003), 305–317.

[9] L. Brugunano and D. Trigiange: *Solving Differential Problems by Multistep Initial and Boundary Value Methods*, Gordon and Breach Science Publishers, Amsterdam, 1998.