# マルコフ連鎖モンテカルロ法

– 統計物理・ベイズ推論以外への応用と手法の改良 –

統計数理研究所 予測制御研究系　伊庭 幸人 [1]

## 1　はじめに

「マルコフ連鎖モンテカルロ法」とは長らく物理学で使われてきた「動的モンテカルロ法」（単に「モンテカルロ法」と呼ばれることもある）に対する情報処理・統計科学の分野での名称である．本稿ではマルコフ連鎖モンテカルロ法の新しい応用及び手法について述べる．当日の講演では以下に述べる内容のうち，１つか２つのトピックに重点を置いて話す予定である．

## 2　応用分野の展開

### 2.1　統計物理，素粒子物理から統計科学へ

マルコフ連鎖モンテカルロ法 (MCMC) の起源が 1950 年代の液体・気体の統計物理に関する研究にあることは良く知られている．その後，この系統のアルゴリズムは，イジング模型など物性論における連続変数・離散変数の確率モデルに応用され，1970 年代末からは相転移・臨界現象の研究に必須の道具となり，また,「格子ゲージ理論」の登場に伴って，素粒子物理における経路積分の数値計算にも利用されるようになった．

短期間のうちに物理学全体で利用されるようになったマルコフ連鎖モンテカルロ法であったが,他の分野への技術移転は遅れた．高次元の非ガウス分布，多変数の離散分布を扱うための強力な手段であるにもかかわらず，そうした要求を共有するはずの統計学・統計的情報処理の分野に本格的に導入されたのはかなり最近である．先駆的なものとして，1970 年代のヘイスティングスによる論文や 1980 年代の尾形・種村による一連の業績はあるものの，統計科学における本格的な登場は 1990 年代を待たねばならなかった．統計科学の分野での「MCMC ブーム」の初期の成果は, 文献 [1] にまとめられている．

なお，工学や統計科学においては,「simulated annealing 法」が先に紹介されたため，MCMC 自体が「最適化の一手法である」（物理関係の人の視点では「物理以外では最適化の手法として使われる」）という偏った理解のされ方をした時期があった．しかし，最近では,「分布からのサンプリングに使うのが本来の姿であり，simulated annealing 法として最適化に用いるのはバリエーション的な用法である」という理解が次第に定着してきているようである．これは情報処理の分野でも確率分布そのものが本質的意味を持つ，という認識と軌を一にしている．

[1]E-mail: iba@ism.ac.jp

## 2.2　新たな応用分野の可能性

90 年代の統計科学の最大のトピックとなった感のある MCMC であるが，MCMC（及び関連手法）が生かせる分野は，統計物理におけるギブス分布のサンプリングと統計科学における（事後）分布のサンプリング以外にはないのだろうか．他にもいろいろな候補が考えられるが，以下ではそのうちの一部を紹介する．

### 2.2.1　時間的経路のサンプリング

興味あるテーマのひとつとして,「確率的力学系の軌道の全体のアンサンブル」を考えてそれを MCMC などでサンプルするという方向がある．量子系のモンテカルロ法で「世界線」をサンプルするというのと概念的には同じであるが，古典系に対して考えることにより，広い視野がひらけてくる．たとえば，分子の解離経路や力学系の閉軌道のような特殊な性質を持つ軌道だけを抜き出してその性質を調べることが可能になる．

「確率過程のパス全体のアンサンブル」という考え方は物理では Onsager に始まるといわれるが，そうしたアンサンブルからモンテカルロ法によってサンプルするというという手法については，1990 年代から，Doll, Zimmer など何人かの研究者による先駆的な仕事 ([2, 3] の文献参照）がある [2]．筆者は以前からこの方向に興味を持って [2]，小規模な計算をいくつか試みてきたが，最近になって D. Chandler たちのグループが精力的に研究をはじめており，分子の準安定状態間の遷移など，化学物理の問題を中心にさまざまな問題にこの方法を適用している [3].

筆者が実際に試してみたことのあるのは,

- イジング模型の “all up” と “all down” の間の遷移経路のアンサンブルからのサンプリング.
- セル・オートマトンの周期解のアンサンブルからのサンプリング.
- ロジスティックマップの周期解のアンサンブルからのサンプリング.
- 「カオス写像の熱力学」の生成関数の計算.

などである．この種の計算は総じて混合（緩和）がきわめて遅くなる上に，カオスに関連する問題では決定論的力学系を（それに雑音を入れた）確率的力学系の極限としてどのように捉えるかという問題があり，いずれも見た目ほど容易ではないことが分かったが，依然として興味を持っている．

### 2.2.2　コンピュータグラフィックス

最近になって，化学物理での研究とは（おそらく）まったく関係なく，コンピュータグラフィクスの分野で，類似の手法が提案されていることを知った．

コンピュータグラフィックスのこの分野では，リアルな動きを人間の手を煩わさずに実現するために,（適当に簡略化・確率化した）運動方程式の解をまじめに計算する．但し，普通の物理系のシミュレーションと違うのは，表現したいことが初期条件の fine tuning を必要とする現象であるこ

[2] 実は，このアンサンブルにクラスター平均場近似などを適用する試みもかなり古くからある．

とである．例としては「ボーリングで特定の２本のピンだけがスプリットで残る」「ビリヤードで何度も玉が衝突して指定された穴に次々に落ちる」などの動画を自動的に生成することを考えればよい．これらの問題は途中の過程を確率過程とし，パスのアンサンブルをモンテカルロサンプリングすることで扱うことができるが，これは化学物理において特定の終状態や遷移状態を通る経路のアンサンブルをサンプルするのと同じである．

かなり特殊な領域であるように思われるが，計算機科学と物理の間の思いがけない関係の例として面白いかもしれない．

### 2.2.3　組み合わせ論

別の興味ある分野としては，さまざまな組み合わせ論の問題がある．完全な数え上げの不可能な問題でも，基底状態（一般に，確率最大の状態）が欲しい配置になっているようなエネルギー（一般に，確率分布）を与えて，モンテカルロ法を適用することで，近似的に解の個数を求めたり，解の特徴を調べることができる．マルコフ連鎖モンテカルロ法など,（物理以外の領域にとって）新しいタイプのモンテカルロ法を用いた例としては，ラテン方格 [4]，魔方陣 [5], N-queen [6] などの研究がある．

### 2.2.4　符号，パッキング

誤り訂正符号の問題にもモンテカルロ法を適用することができる．ひとつは，符号の解読そのものに利用することである．この場合，受信した符号をデータとみなして，事後分布からのサンプルを行うわけであるから，本質的にはベイズ統計への応用と同じともいえるが，符号の種類によってはかなり違う面もあると思われる．モンテカルロ法の使用を前提として新しい符号を考えたり，その事後分布の構造を調べたりすることといった方向も興味深い．

別の応用としては，符号の符号語の生成にモンテカルロ法を利用することも考えられる．この場合，たとえば，高次元ビット空間への符号語のパッキングの問題を考えることになる．統計数理研究所の伊藤 [7] はランダム充填による Goley 符号の生成を論じているが，このような問題を分布からの正確なサンプリングを可能にするようなモンテカルロ法で扱うことで，より定量的にパッキングの構造を知ることができるかもしれない．

## 3　手法の拡張

これまで述べた分野の問題には，マルコフ連鎖モンテカルロ法にとって，非常に厳しい条件のものが多い．状態空間の landscape は multi-valley 構造やぎざぎざの多い構造を示し，これらは著しく混合（緩和）を阻害する．従って，これらの問題を扱うためには，従来にも増して，手法の改良の努力が必要となる．ここでは，最近有力となっている２種類の方向について簡単に述べる．

### 3.1 拡張アンサンブル法

ひとつの方向としては，もとの分布を拡張して，あるいは分布族を合併して扱うことで，効率の良い計算を可能にする手法群がある．マルチカノニカル法，アンブレラ法，パラレル・テンパリング法などがこのカテゴリーに属する．詳しくは筆者による総合報告 [8] を参照されたい．

### 3.2 Population 型のモンテカルロ法の再発見と拡張

長い歴史を持つが，最近再び脚光を浴びている考え方として，多数の系 (particle, walker) を並列にシミュレートして，その過程でそれぞれに適切な重みを定義し，重みの大きいものを分裂させ，小さいものは消すことで importance sampling を実現するというものがある．

最も良く知られた例は Kalos らによって導入された「拡散モンテカルロ」「グリーン関数モンテカルロ」である．これらは歴史的に量子系を対象としているが，原理的には古典的な確率計算にも利用でき,「転送行列モンテカルロ」などの名前で知られている．1990 年代になって，このタイプの手法が統計科学と高分子物理で独立にリバイバルした．前者は「パーティクルフィルタ」「逐次モンテカルロ」などの名で時系列解析など各種の統計的推論の問題で有効性が示されている．後者は「PERM」の名でポリマーや格子タンパクの折りたたみの問題に応用されている．これらの方法は，最適化手法としての遺伝的アルゴリズム（GA) に制限を加えて，確率分布に関する計算を厳密に行えるようにしたものとも考えられる．

これらの研究はいくつかの領域で独立に行われていたようであるが，筆者の総合報告 [9] が，それらの分野を包括的に扱い，分野間の関係を明らかにした最初のものと思われる．

なお.「逐次モンテカルロ」も「PERM」も時系列の時間軸方向やポリマーの成長する方向にシミュレーションを進めていくが,「拡散モンテカルロ」等でもわかるように，他のパラメータが変化する方向に変化させることもできる．そうした場合も含めて考えることで，simulated annealing やマルコフ連鎖モンテカルロ法との関係がより良く理解できる． すでにこうした方向の研究もいくつかあるが，今後，重要な方向と思われる．

## 参考文献


[1] W. R. Gilks, S. Richardson and D.J. Spiegelhalter (ed.), *Markov Chain Monte Carlo in Practice*, (Chapman and Hall, London, 1996).

[2] 伊庭幸人，無時間の思想，基研研究会「認知・行動の基底としての力学と論理」報告,「物性研究」**71-2** (2000 年 11 月号).

[3] Chandler Group の論文一覧：
`http://gold.cchem.berkeley.edu:8080/bibliography.html`
上記のリストで，論文番号 150, 152, 156, 158, 165, 167, 172, 176, 180, 184，. . . 等を参照.

[4] M. T. Jacobson and P. Matthews, Journal of Combinatorial Designs, **4** (1996) 405.

[5] K. Pinn, Int. J. Mod. Phys. C **9** (1998) 541.

[6] K. Hukushima: Comp. Phys. Comm. **147** (2002) 77.

[7] Y. Ito, Ann. Inst. Statist. Math., **38** (1986), 583.

[8] Y. Iba, Extended Ensemble Monte Carlo, Int. J. Mod. Phys. C **12** (2001) 623.

[9] Y. Iba. Population Monte Carlo algorithms, Transactions of the Japanese Society for Artificial Intelligence **16** No.2, (2001) 279, `cond-mat/0008226`.