科研費特定領域研究「確率的情報処理への統計力学的アプローチ」主催
日本神経回路学会，電子情報通信学会基礎・境界ソサイエティ共催
若手研究者・学生向けに最新技術をわかりやすく紹介する講演会「確率的アルゴリズムによる情報処理」

モンテカルロ法の前線

— サイコロ振って積分する方法 —

東京大学 大学院総合文化研究科 広域科学専攻 相関基礎科学系　福島 孝治 [1]

1　はじめに

モンテカルロ法 (Monte Carlo) は 1950 年代の初期に液体のシミュレーション方法として導入された [1]．その後，主に統計物理の分野 [2] において，様々な問題に応用される一方で，多くの改良，あるいはその亜種を産み出してきた．1980 年代に統計科学・計算科学の分野に，マルコフ連鎖モンテカルロ法 (Markov chain Monte Carlo, MCMC) やアニーリング法 (Simulated Annealing) として，応用範囲を広げていった．三省堂「大辞林」によれば，モンテカルロ法とは，

> 「偶然現象の経過をシミュレーションする場合に、乱数を用いて数値計算を行い、問題の近似解を得る方法。コンピューターの発達によって広い分野で利用されている。」

のようであるが，中らずとも[2]遠からずと言える．現在では，モンテカルロ法は，乱数を用いた一般のシミュレーション方法の総称と考えられている．本稿では，その中で特に狭義のモンテカルロ法として，マルコフ連鎖モンテカルロ法と最近の発展について解説する．

ここでのモンテカルロ法の目的は，d 次元空間で点$\boldsymbol{X}$(成分を $x_i, i = 1, \cdots, d$ とする) の従う確率分布 $P(\boldsymbol{X})$ が与えられたときに，

1. 確率分布 $P(\boldsymbol{X})$ に従うように点 $\{\boldsymbol{X}^{(i)}\}$ $(i = 1, \cdots, M)$ をサンプリングする，
2. または，その分布についてのいろいろな関数の期待値

$$\langle A \rangle = \int_D A(\boldsymbol{X}) P(\boldsymbol{X}) d\boldsymbol{X}, \tag{1}$$

を計算する

ことである．対象となる分布関数として，統計物理のギブス分布，またはベイズ統計での事後分布などがその一例であるが，一般に，大自由度の確率モデルに関する高次元積分法と考えれば，その対象はさらに広がることがわかるであろう．特に，変数$\boldsymbol{X}$ の次元 d が非常に大きいときにモンテカルロ法は有効であると考えられている．

2　モンテカルロ積分から

モンテカルロ法と聞くと，乱数を用いた円周率 π の計算を思い浮かべる方が多いのではないだろうか．WEB で「モンテカルロ法」を検索してみても，軒並「π の計算」が紹介されている．モ

[1] E-mail: hukusima@phys.c.u-tokyo.ac.jp, URL: http://www.dbs.c.u-tokyo.ac.jp/~fukushima/
[2] 統計的な誤差を含むという意味では近似解という言い方は正しいが，系統的な誤差はないことが多い．

ンテカルロ法の旨味は，必ずしも単純なモンテカルロ積分では説明しきれないが，モンテカルロ法の基本要素の一部はそこに含まれている．まずは，モンテカルロ積分から見て行くことにする．

ある高次元空間で定義される関数 $g(\boldsymbol{x})$ の積分領域 D についての多重積分 $I = \int_D d\boldsymbol{x} g(\boldsymbol{x})$ は，確率変数 x の確率分布を $P(X)$ としたとき，期待値の問題

$$I = \int_D d\boldsymbol{x} g(\boldsymbol{x}) = \int_D d\boldsymbol{x} \frac{g(\boldsymbol{x})}{P(\boldsymbol{x})} P(\boldsymbol{x}) = \left\langle \frac{g(\boldsymbol{x})}{P(\boldsymbol{x})} \right\rangle \tag{2}$$

に書き直すことができる．分布 $P(X)$ に従うように変数 x を抽出 (サンプリング) できれば，期待値として積分が計算できることになる．$P(X)$ を一様分布とする単純モンテカルロ積分では，積分領域 D から一様にランダム・サンプルした M 個の代表点，$\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(M)}$ で積分 I

$$\widetilde{I} = \frac{D}{M} \left\{ g(\boldsymbol{x}^{(1)}) + g(\boldsymbol{x}^{(2)}) + \cdots + g(\boldsymbol{x}^{(M)}) \right\}. \tag{3}$$

を近似的に求めるわけである．大数の法則により，代表点の個数 M の大きな極限で近似値 $\widetilde{I}$ は I になることがわかる．この場合，モンテカルロ積分の誤差は，$O(M^{-1/2})$ と見積もることができる．例えば，一次元積分として，$g(x) = 4\sqrt{1-x^2}$ とすれば，領域 $x \in [0,1]$ での積分は π の計算[3]になる．ところが，普通は単純モンテカルロ積分で π を評価することは絶対にしないし，日本が誇る π の世界記録もモンテカルロ積分では不可能である．なぜならば，誤差が非常に大きいからである．積分領域を M 個の点で均一に区切った最も単純な積分の台形公式でさえ，その真値からのずれは $O(M^{-2})$ であるので，モンテカルロ積分の誤差 $O(M^{-1/2})$ はとても悪く思える．

しかしながら，モンテカルロ積分の誤差の評価は被積分関数の変数の次元に原理的には依存しない．一方で，一般の求積法では次元が大きくなるほどその手続きは増える．d 次元積分領域を均一区間で区切るためには，$O(M^d)$ の手間がかかるわけである．このことから，非常に高次元のある種の積分を行いたい場合には，モンテカルロ法による評価が有効になることがわかる．

3　メトロポリスの方法へ

ただし，一様なランダム・サンプリングを用いた単純モンテカルロ積分を高次元積分に応用するのは，以下の困難点がある；(1) 高次元空間の中で対象となる積分領域を一様にサンプルすることは意外に難しく，(2) ほとんどの点で被積分関数の値が小さく，無駄が多く (図 1)，結果として評価した分散が極めて大きくなってしまう[4]．

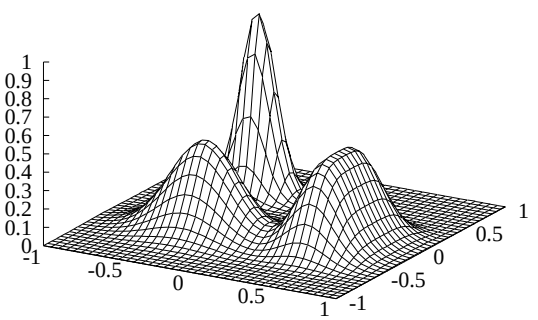

図 1: 二次元分布関数 $f(x,y)$ の例．二変数でさえ関数値が小さい領域が多くあり，ランダムに選んだ点 (x,y) が有為な値をもつ確率は低そうである．さらに $10^4 \sim 10^5$ 程度の高次元分布関数になるとさらにその傾向は大きい．

[3]もっとも，よく知られている π の計算は二次元積分を考え，長さが 1 の正方形上のランダムな点が単位円の中に位置する割合を求める方法が多い．これは棄却サンプリング法と呼ばるモンテカルロ法の一種である．

[4]棄却サンプリング法でも高次元積分では似たような困難がある．

この困難を解決したのが，今日メトロポリス法として多くのモンテカルロ法の雛型になっている方法である．この方法では，それぞれの成分を一様にランダム・サンプルするのではなく，積分に重要な寄与のあるところを重点的にサンプルしようという考えに基づいている．具体的には，対象となる分布関数 $P(X)$ に比例するようにサンプルを行うのである．もしもこれが可能になれば，サンプル点 $X^{(1)}, X^{(2)}, \cdots, X^{(M)}$ から期待値 (1) は，

$$\widetilde{\langle A \rangle}_T = \frac{1}{M}\left(A(X^{(1)}) + A(X^{(2)}) + \cdots + A(X^{(M)})\right), \tag{4}$$

として求めることができる．

それではどのようにサンプルすればいいのであろうか．確率分布 $P(X)$ が具体的にわかっていれば，逆関数を求めることにより，目的は達成されるであろう．しかし，一般に多変数関数の逆関数を求めることは難しく，そもそも $P(X)$ の関数形はわかっていない．メトロポリス法の基本的な考え方は，確率過程を用いて逐次的に点 X を生成することである．この確率過程では現在の点 $X^{(i)}$ は直前の点 $X^{(i-1)}$ にのみ依存した確率分布に従っており，これ故に，この点列 $\{X^{(i)}\}$ はマルコフ連鎖と呼ばれる． マルコフ連鎖で作られる点の集合 $\{X\}$ が対象となる確率分布 $P(X)$ に従うようにするのが我々の要請である．つまり，構成したい分布関数に収束するように，点 X から X' への遷移確率 $W(X \to X')$ をうまく決定したいわけである．W が確率であるための条件として，任意の点 X, X' に対して，$0 \leq W(X \to X') \leq 1$ であり，保存則 $\int dX W = 1$ を満たさなくてはならない．これらの条件から，この確率過程の定常分布の存在が保証される．また，その定常分布が求めたい確率分布 $P(X)$ になるための必要条件は以下の 2 つである．

(A) 詳細つりあいの条件 ： 遷移確率は関係式

$$P(X)W(X \to X') = P(X')W(X' \to X), \tag{5}$$

を満たす．

(B) エルゴード条件： 任意の 2 つ点 X と X' の間の遷移確率がゼロではないか，あるいは有限個のゼロでない遷移確率の積で表される．

条件 (A) より確率分布が不変・定常分布になることがわかる．また，条件 (B) は空間のどんな 2 点も有限の遷移確率で連結していることを意味しており，これらの条件から任意の初期状態 $X^{(0)}$ からつくられるマルコフ連鎖が唯一の定常分布 $P(X)$ に収束することが示される．

この条件だけから遷移確率は一意的には決定できないが，メトロポリス型と呼ばれる例として,

$$W(X \to X') = \min\left(1, \frac{P(X')}{P(X)}\right) \tag{6}$$

は式 (5) を満たしている．このような遷移確率で得られたマルコフ連鎖から，十分な間隔で抜きだした状態点の集合 $\{X^{(i)}\}$ を求めたい確率分布 $P(X)$ からの独立なサンプルとみなし，興味ある量 $A(X)$ の確率分布 $P(X)$ に関する期待値等を計算する．

メトロポリスのモンテカルロ法をアルゴリズムとしてまとめておく．

ステップ **0**：初期設定 任意の初期状態 $X^{(0)}$ を選ぶ．

ステップ **1**：次の状態の候補 現在の状態 $X^{(t)}$ からランダムに変化させた新しい状態 X' を作る．

ステップ **2**：遷移確率の計算 遷移コスト $P(X')/P(X^{(t)})$ を計算する．

ステップ **3:** 状態更新 一様乱数 $r \in [0,1]$ を生成し，以下のように次の状態 $X^{(t+1)}$ を決める;

$$X^{(t+1)} = \begin{cases} X' & \text{if } r \leq P(X')/P(X^{(t)}) \\ X^{(t)} & \text{otherwise} \end{cases} \tag{7}$$

ステップ **4**：くりかえし ステップ 1 へ戻る．ステップ 1 から 4 までの試行を 1 モンテカルロステップと呼ぶことにする．

このメトロポリス法は，当初は統計物理のシミュレーション法として提案されたが，ここまでのところ，物理の特殊性は何も使っていないことがわかる．確率分布がギブス分布であることは方法の構成のためには全く関係ないのである．この点がモンテカルロ法の汎用性を表している．

3.1 熱浴法，ギブスサンプラー

先に述べたように，遷移確率 W は他にも選択の自由度はあり，たびたび用いられるのが熱浴法 (ギブスサンプラー) である．簡単のために状態 X の各成分 x_i は離散的な値 $\{\alpha_i^\mu, \mu = 1, \cdots, A_i\}$ しかとらないとする．現在の状態 $X = (x_1, \cdots, x_d)$ から i 成分を更新するときに，今の状態も含めて A_i 通りの可能性があるが，それぞれの状態 $X_\mu = (x_1, \cdots, \alpha_i^\mu, \cdots, x_d)$ への遷移確率は，

$$W(X \to X_\mu) = \frac{P(X_\mu)}{\sum_{\mu'} P(X_{\mu'})} \tag{8}$$

で与えられる．x_i が連続変数の場合は分母の和は積分となる．この方法では，メトロポリス法と異なり，W は現在の状態には依存しない．どちらがよいかを判断するはっきりとした決め手はない．

3.2 統計物理の例題：イジング模型の場合

具体的な例として，統計物理での磁石のモデルであるイジング模型を取り上げる．統計力学では，対象とする物理系のミクロな状態 X のエネルギーを $E(X)$ としたときに，確率分布はギブス分布

$$P(X;T) = \frac{1}{Z(T)} \exp(-E(X)/T) \tag{9}$$

で与えられ，この分布関数での (熱平衡) 期待値に興味がある．ここで T は温度と呼ばれる量であり，$Z(T)$ は分布関数 $P(X)$ の規格化定数である．統計物理の対象は様々であるが，対象をモデル化する操作はミクロな状態 X の指定とそのエネルギー関数 $E(X)$ を具体的に与えることに相当する．三次元立方格子上のイジング模型では，各格子点 i に ± 1 の二値を取るスピン変数 S_i，すなわちミクロな磁石を置く．スピン変数の値 ± 1 はそれぞれ「上向き」と「下向き」を表している．一辺の長さを L とすると $N = L^3$ 個のスピン変数がミクロな物理状態 $X = \{S_1, S_2, \cdots, S_N\}$ を指定する．状態のとりうる領域，すなわち積分領域は $X \in [-1,1]^N$ であり，厳密に積分式 (1)(今の場合は多重和) を評価するのは N が 30 程度でほぼ不可能となる．

今，近接するスピン変数は同じ方向に揃いたがる系 (強磁性模型) を考えると，エネルギー関数は，

$$E(\{S\}) = -\sum_{\langle ij \rangle} J S_i S_j, \tag{10}$$

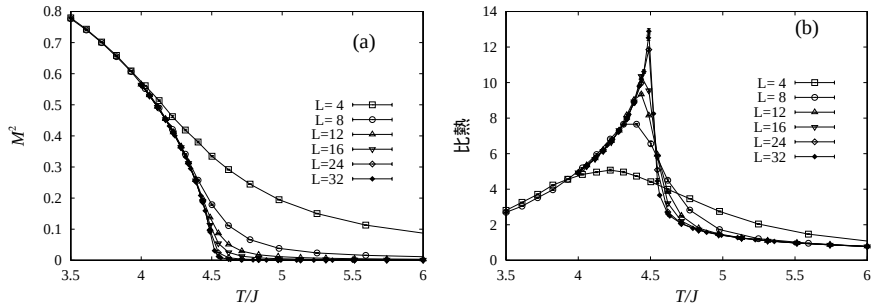

図 2: 三次元イジング模型の二乗磁化 (左) と比熱 (右) の温度依存性．複数のサイズ N についての結果が描かれている．比熱は温度 $T/J \sim 4.5$ あたりで発散の傾向を示し，磁化はその温度より低温有限の値を持っている．これらは強磁性相転移を示唆している．

で与えられる．和 $\sum_{\langle ij \rangle}$ は隣接する格子対についてとることにする．このような三次元格子上の模型で厳密 な解答は得られておらず，磁石の強さを表す二乗磁化 $\left(\frac{1}{N}\sum_i S_i\right)^2$ のような量のギブス分布に関する期待値を温度 T の関数として求めたいわけである．

それぞれのステップでの具体的な操作は以下のとおりである．

ステップ **0**：初期スピン配位 各格子点上にスピン変数 S_i を確率 $1/2$ で 1 か -1 を与える．この例での状態 X はスピン変数の集合 $X^{(0)} = \{S_1^{(0)}, S_2^{(0)}, \cdots S_N^{(0)}\} = \{1, -1, \cdots, -1\}$ として表される．ここで，$S_i^{(t)}$ は，t モンテカルロステップ時の格子点 i のスピン変数の値を意味する．

ステップ **1**： 次のスピン配位の選択 (メトロポリス法) ランダムに選んだ格子点 k のスピンを反転したスピン配位を次の状態の候補とする．$X' = \{S_1^{(t)}, S_2^{(t)}, \cdots, -S_k^{(t)}, \cdots, S_N^{(t)}\}$.

ステップ **2:** 遷移確率の計算 遷移コスト Δ を評価する．

$$\begin{aligned} \Delta &= \frac{P\left(X'\right)}{P\left(X^{(t)}\right)} = \frac{\exp\left(-E(\{S'\})/T\right)}{\exp\left(-E(\{S^t\})/T\right)} = \frac{\exp\left(\sum_i{}' JS_i^{(t)}(-S_k^{(t)})/T\right)}{\exp\left(\sum_i{}' JS_i^{(t)}S_k^{(t)}/T\right)} \\ &= \exp\left(-2\sum_i{}' JS_i^{(t)}S_k^{(t)}/T\right) \end{aligned} \tag{11}$$

ここで，対象とする確率分布がギブス分布 (9) であり，イジング模型のエネルギー関数が式 (10) であることを使った．最後の結果は，現在のスピン配位の注目する格子点 k の近接の和だけで表されている．これは，一つのスピン反転に伴う計算手続きが系の大きさと無関係に $O(1)$ で可能なことを意味している．

ステップ **3**：乱数との比較 一様乱数と比較することにより，ルール (7) に従って，次の状態 $X^{(t+1)}$ を決定する．一つ注意する点は，この更新で実際の状態の変化が無い場合 $(X^{(t+1)} = X^{(t)})$ でも，状態更新と見なして，物理量の平均には考慮する必要がある．

ステップ **4**：くりかえし ステップ 1 に戻って，同じ操作を繰り返す．この操作からわかるとおり，メトロポリス法で作られるサンプル列 $\{X^{(1)}, X^{(2)}, \cdots, X^{(M)}\}$ の間には強い相関が残ることが推測される．この点にメトロポリス法と単純なモンテカルロ積分との顕著な違いがあり，ある種の困難さが含まれている．この問題は次章に詳しく議論する．

このように構成されたモンテカルロ法による実際の計算結果を示しておく．図 2 は，三次元イジング模型 (10) の様々な大きさ ($N = 4^3 \sim 32^3$) の系についての二乗磁化と比熱 $\frac{1}{T}\left(\langle E^2\rangle - \langle E\rangle^2\right)$ の温度依存性の結果である．全スピン数 N を大きくするにつれて，ある温度 ($T_c/J \sim 4.51$，転移温度) を境に，磁化はゼロから有限の値へ定性的な変化をしていることがわかり，その転移温度で比熱は発散の傾向を示している．この転移温度周辺の特異性は相転移の臨界現象と呼ばれる．ここでの最大の $N = 32^3$ では，近似している和の総数は 2^{32768} となっている．モンテカルロ法以外では無理な数字であろう．

4　定常分布への遅い緩和の問題

先にモンテカルロ法の方法論が非常に一般的であることが応用範囲を広くしていると述べた．実際に沢山の成果を挙げている [2, 7]．しかしながら，同時に様々な系への適用を通して，ある共通の困難に遭遇することがわかった．それは，ある条件のもとでマルコフ連鎖において状態の遷移確率 $W(X \to X')$ が小さくなり，実質的に遷移が行なわれなくなることである．結果として，独立とみなせるサンプル数が極めて少く，非常に効率が悪くなることである．また，初期条件に強く依存し，間違った答えを導くこともある．広大な高次元空間の中で，サンプリングされる点がある局所的な部分空間に留まり続けているという意味で，この現象は遅い緩和の問題と呼ばれている．このために，確率分布からの正しいサンプリングができず，必要な精度が得られないことが起こり，モンテカルロ法の応用範囲は狭められている．そして，この問題は幾つかの興味深い状況で現れるのである．それは，物理の問題では相転移温度近傍であり，または競合する複雑な相互作用により確率分布が多峰的になる場合である．

後者の物理系を一つ挙げることにする．前節で簡単なイジング模型を紹介し，そのモンテカルロ法による結果を示した．もちろん，温度が $4.51J$ あたりで，スピンが揃い始めることはすぐにはわからないが，非常に低温ではエネルギー関数のより小さな値をとる状態 X，すなわち，すべてのスピンが揃っている状態付近の確率が大きくなる．しかし，エネルギー関数 (10) の中の相互作用 J を格子点 ij に依存して，ランダムに ± 1 の値にするモデルでは状況は完全に変わってくる．図 3 に二次元正方格子の場合のある相互作用 $\{J_{ij}\}$ での最もエネルギーが低い状態の例を示した．もはやスピンが奇麗に揃った状態はエネルギー的には低い状態ではなく，むしろ相互作用で決まるランダムな配位の方がギブス分布は大きな値となる．低温ではスピンがガラスのようにランダムに凍結することが期待されるので，この模型はスピングラスと呼ばれている．この模型の重要な性質の一つは，最低エネルギー状態あるいは，それに近い状態が非常に多数，しかも空間の離れたところ

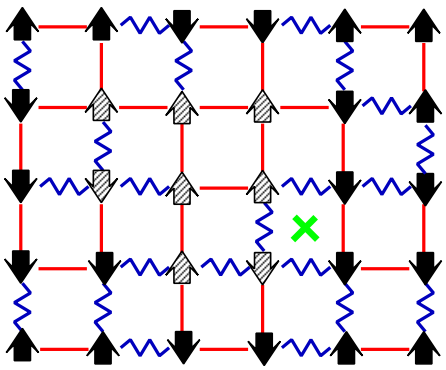

図 3: 二次元正方格子上のスピングラスの例． 実線 (相互作用 J_{ij} が正) で継れたスピンは平行に，波線 (J_{ij} が負 0) では反平行になるとエネルギー関数は小さくなる．このために× マークを囲む四角形の回りのエネルギーを全て同時に低くするスピン配位は存在しない (このような四角形の数は全部で十個あるが，おわかりだろうか)．図のスピン配置は最もエネルギーが低い状態であるが，他にも同じエネルギー状態が多数存在する．例えば，斜線スピンを同時に反転させても最低エネルギー状態となる．対応する三次元スピングラス模型の最低エネルギ状態を探す問題は NP 問題である．

に存在することである．結果として，ギブス分布は沢山のピークを持つ．イメージとしては，図 1 のような多峰的な分布が実現する．このような系にモンテカルロ法を適用すると，遷移確率を使ってあるピークから別のピークへ状態を変化させる必要がある．少なくとも一度は分布関数の値の小さな領域を通過しなければならない．そのようなピークからピークへの遷移が非常に希にしか実現しないのが，難しさの本質的なところである．

確率的に最適な状態に対応する峰の他に，沢山の準最適な峰が存在するような状況はスピングラス以外にも，タンパク質や高分子等の複雑なモデルでも起こっている．どうやら，問題に多数の制約条件があることが鍵になっているようであり，最適化問題の困難さの一つも同様なことと考えられる．次章では，最適化手法からこの困難点を解消する術を議論する．

5　アニーリング法は最適化技法

モンテカルロ法の最適化問題への展開は有名なシミュレーテッド・アニーリング法である．アニーリング法はその名前からわかるとおり物理現象の焼きなましが起原であり，実際に物理学者により提案されている [3]．この方法は，モンテカルロ法と同様に遷移確率 (6) を用いてマルコフ連鎖を構成しようとするが，定常分布が実現するまでに分布関数 $P(X)$ に含まれる問題の難しさを表すパラメータを徐々に変化させることが特徴である．ギブス分布では，温度がちょうどそのパラメータであるが，一般に考えている分布関数 $P(X)$ に対して温度パラメータ T を導入し，

$$P(X;T) = \frac{P^{1/T}(X)}{\sum_X P^{1/T}(X)} \tag{12}$$

のようなパラメータ変形した分布を考える．高温極限 $T=\infty$ では分布 $P(X;T)$ は元の分布 $P(X)$ に依らず，一様分布であり，マルコフ連鎖で更新される点 X は状態空間でランダムウォークをする．分布関数 $P_{T=\infty}$ はピークを持たないので，ピーク間の遷移が困難になることはありえない．一方，低温では，分布 $P(X)$ を最大にするようなエネルギーの低い状態が実現される．アニーリング法の要点は，高温での探索能力を生かしながら，徐々に温度を下げ (焼きなまし) ながら，絶対零度で最適解を見つけようとすることにある．

最適化手法としてのポイントは，確率的な遷移の導入により，必ずしも最適解に近付かない場合，すなわちエネルギーが上がってしまう場合も状態遷移を許していることであり，パラメータ T の変化の際に詳細つりあいの条件を破っていることがモンテカルロ法と異なる点である．この意味でアニーリング法はあくまで最適解探索用に構成された最適化手法方法であり，分布に比例するようなサンプリングができる保証はない．アニーリングの途中の温度で現れる状態の集合から統計平均を評価しても，確率分布からの期待値にはならない．特に遅い緩和を示す問題では，正しい期待値からのずれは顕著に現れる．

しかしながら，アニーリング法では高次元空間の分布に多峰性があるときに，複数の峰からその重みに従うようにサンプリングを行うための指針を与えてくれているように思える．では，どのようにすればよいであろうか？例えば，高温から温度を下げて，ある低温で固定してマルコフ連鎖を続けるとする．その場合は，状態空間の局所的な探索により求めたい分布の一部は再現できるが，あるピークからの脱出や新しい状態の探索能力は劣り，分布関数全体の様子は分からなくなってしまう．一方，温度を人為的に上げたり下げたりすることで，ピーク間の遷移を図れば，今度は詳細つりあいの条件を壊してしまい，分布の局所的な性質やピーク間の相対的な重みの情報を失って

しまう．詳細釣り合いによる分布の不変性とピーク間の有効な遷移による多様性の獲得の間で，アニーリング法はジレンマに陥ってしまっているのである．

6　拡張アンサンブル法

前節のジレンマを解消し，遅い緩和問題を解消するために，マルチカノニカル法 [4] や交換法 [5] 等の一連の方法群が近年提唱された [6, 7]．これらの方法は，興味ある分布関数に修正を加えたり，合体させたりした拡張された分布関数を取り扱うことから「拡張アンサンブル法」と呼ばれている [6]．この考え方は，アニーリング法での分布の変形 (12) に近く，更にサンプリング手法として欠かせない詳細つりあいを満たしていることが重要な点である．拡張アンサンブル法は分布に多数の峰を持つような系に対しても有効であることが示されてきている [8]．ここでは，幾つかの技法について紹介する．

6.1　交換法 [5]

交換法では，Q 個の異なる温度 $(T_1, T_2, \cdots, T_Q)$ に接触しているレプリカ系を用意する．このパラメータは必ずしも温度である必要はなく，対象となる確率分布を特徴付ける変数であればよい．以下ではギブス分布の温度の場合について考える．レプリカの指標を q として，状態を $\{X\} = \{X_1, X_2, \cdots, X_Q\}$ と表したとき，拡張された同時分布関数 $P_{\rm EX}(\{X\})$ は，

$$P_{\rm EX}\left(\{X\}\right) = \prod_{q=1}^{Q} P_q(X_q) \tag{13}$$

と表される．$P_q(X)$ は温度 T_q で特徴づけられる確率分布である．交換法のマルコフ連鎖は，この同時分布関数を不変にするように 2 種類の状態更新を導入する．一つは，それぞれのレプリカ系の分布関数 $P_q(X)$ に対する式 (6,7) のような通常の状態更新である．もう一つは，レプリカ間の交換である．すなわち，2 つの温度 T_q と T_{q+1} の状態 X_q と X_{q+1} を交換するのである．式 (13) を不変にするためには，交換レート R を

$$R = \frac{P_{\rm EX}(X_1, X_2, \cdots, X_{q+1}, X_q, \cdots, X_Q)}{P_{\rm EX}(X_1, X_2, \cdots, X_q, X_{q+1}, \cdots, X_Q)} = \frac{P_q(X_{q+1}) P_{q+1}(X_q)}{P_q(X_q) P_{q+1}(X_{q+1})} \tag{14}$$

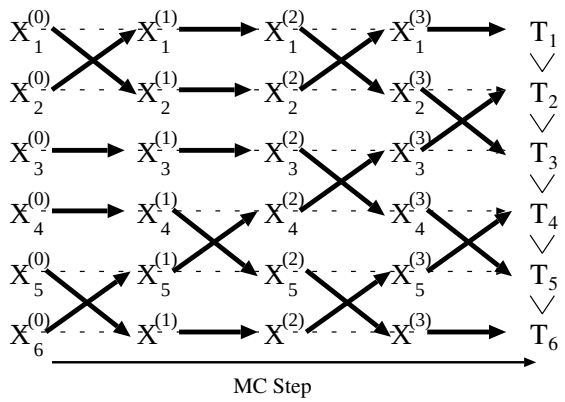

図 4: $Q = 6$ レプリカ系での交換法の状態更新の様子．交換のプロセスを通じて，状態は様々なパラメータ (温度) の確率分布を経巡る．

とし，メトロポリス型ならば交換遷移確率を $\min(1, R)$ とすればよい．ギブス分布の交換レートは，

$$R = \exp\left(\frac{1}{T_q} - \frac{1}{T_{q+1}}\right)(E(X_q) - E(X_{q+1})), \tag{15}$$

となり，交換すべき 2 つのレプリカの温度差とエネルギー差から計算できる．これらの 2 つの過程により，各レプリカの状態はそれぞれの温度で更新をしながら，同時に温度軸上を動き回り，長時間の振舞はランダムウォークとみなすことができる．つまり，このマルコフ連鎖では詳細つりあい条件を満たしながら，自分自身でアニーリング (とヒーティング) を行っているように見える (図 4)．結果として，低温であるピークに留まり続けることがあったとしても，一度高温になることにより簡単に脱出できることを期待するわけである．ある温度 T_q での期待値は，通常のように $P_q(X_q)$ の平均として計算できる．実は，図 2 の結果は交換法によって得られた結果である．

6.2 マルチカノニカル法 [4]

交換法では複数の確率分布からなる同時分布関数を考えたが，マルチカノニカル法はエネルギーに対する修正された確率分布を用いる．調べる系の状態密度 $D(E)$ をエネルギーが $E < E(X) < E+dE$ の範囲にある状態 X の数と定義すると，マイクロカノニカル分布は，

$$P_{\rm micro}(X; E) = \frac{\delta(E(X) - E)}{D(E)} \tag{16}$$

で与えられる．マルチカノニカル法のポイントは，ギブス分布に比例するように状態 X を発生させるかわりに，適当なエネルギーの範囲 $E \in [E_{\rm min} : E_{\rm max}]$ のマイクロカノニカル分布を同じ重みで重ね合わせた混合分布

$$P_{\rm mul}(X) \equiv \int_{E_{\rm min}}^{\rm E_{max}} dE P_{\rm micro}(X; E) \tag{17}$$

を不変分布とすることである．これは単純に $P_{\rm mul}(X) \sim 1/D(E(X))$ となる．その結果，エネルギーの周辺分布は E に依らない定数になり，系はエネルギー軸上をランダムウォークすると考えられる．このことから，低エネルギーでの準安定状態からの脱出を期待できる．これは交換法が温度空間をランダムウォークすることに似ているが，マルチカノニカル法の場合は状態空間を増やしたのではなく，エネルギー分布を変更していることに注意されたい．

マルコフ連鎖の構成には，詳細つりあいの式 (5) において，$P(X) = D^{-1}(E(X))$ として，メトロポリス型ならば遷移確率 (6) に従って状態を更新する．実はシミュレーションの前に状態密度 $D(E)$ なる量はわからないわけだが，必ずしも厳密な $D(E)$ を知る必要はないので，短いシミュレーションを繰り返しながら順次 $D(E)$ の関数形を学習していく方法が使われている．任意の関数 $A(X)$ のギブス分布に関する期待値は，長い観測用シミュレーションから，再重率法

$$\langle A \rangle_T = \frac{\sum_i A(X^{(i)}) \exp(-E(X^{(i)})/T) P^{-1}(X^{(i)})}{\sum_i \exp(-E(X^{(i)})/T) P^{-1}(X^{(i)})} \tag{18}$$

により様々な温度パラメータで一度に計算できる．

7 おわりに

本稿ではモンテカルロ法の初歩から最近の進展までを概観した．今年は，1953 年にメトロポリスらによるモンテカルロ法の提唱 [1] から，ちょうど 50 周年を迎える．それを記念して，発祥の地

であるロスアラモス国立研究所においてワークショップが開かれ[5]，方法論としての発展が多いに議論された．ここで触れられなかった技法はまだまだ沢山ある．特に，マルコフ連鎖モンテカルロ法と双璧をなす逐次モンテカルロ法，あるいは広くポピュレーション型モンテカルロ法と呼ばれる方法群は近年よく応用されている．興味のある方は文献 [7, 9, 10] を参照頂きたい．

また，一方，方法論の発展のすぐ傍で多数の応用例が統計物理あるいは素粒子物理の一部，またベイズ統計へ展開され，大きな成果を挙げ続けてきた．特に，拡張アンサンブル法は複雑な系を舞台に応用され，その威力を発揮しはじめているところである．本稿で，そのような各分野での成果には紹介できず，実際の有効性について雰囲気すら伝えられていないのは，筆者の力量不足に他ならない．是非，最近のレビュー [6, 8] や文献 [7] を参照していただきたい．また，読者がそれぞれの興味ある対象でモンテカルロ法を実践され，その有効性を体験されるならば，それは筆者の望外の幸せである．

謝辞

本稿の話題については，統計数理研究所の伊庭氏との共同研究や日頃の議論から学ぶところが非常に大きい．ここにあらためて感謝の意を表したい．

参考文献

[1] N. Metropolis, et. al. : Equations of state calculations by fast computing machines, J. Chem. Phys. **21**, 1087/1091 (1953)

[2] D. P. Landau and K. Binder: A guide to Monte Carlo simulations in statistical physics, Cambridge University Press (2000).

[3] S. Kirkpatrick, C.D.Gelatt, and M.P.Vecchi, “Optimization by simulated annealing”, Science **220** 67/680, 1983.

[4] B. A. Berg and T.Neuhaus: Phys. Lett. **B267** (1991) 249; Phys. Rev. Lett. **68**(1992)9.

[5] K. Hukushima and K. Nemoto: Exchange Monte Carlo Method and Application to Spin Glass Simulations, J. Phys. Soc. Jpn., **65**, 1604/1611, (1996).

[6] Y. Iba: Extended Ensemble Monte Carlo, Int. J. Mod. Phys. C, **12**, 623/652, (2001).

[7] J. S. Liu: Monte Carlo Strategies in Scientific Computing, Springer-Verlag New York (2001).

[8] 例えば，タンパク質系への応用は，杉田 有治，光武 亜代理，岡本 祐幸: “拡張アンサンブル法によるタンパク質の折り畳みシミュレーション”, 日本物理学会誌 Vol.**56**, pp591, (2001).

[9] A. Doucet et al. (eds.), *Sequential Monte Carlo in Practice*, Springer-Verlag (2001).

[10] Y. Iba, Trans. Jpn. Soc. Artif. Intel. **16**, 279 (2001), e-print: cond-mat/0008226.

[5] http://cnls.lanl.gov/Conferences/MonteCarloMethods/