

12.5 モンテカルロシミュレーションの技法

12.5.1 原理

モンテカルロシミュレーションとは一般に乱数を用いる数値シミュレーションの総称である。本節では、その中でも与えられた温度での熱平衡状態を作り出す方法としての狭義のモンテカルロ法(メトロポリス法)を、特に神経回路網への応用を念頭において解説する。もちろん、神経回路網以外のシステムへの拡張も容易にできる。なお、神経回路網や情報処理へのモンテカルロ法の重要な応用として、温度を順次下げてゆくことにより最適解を探索するシミュレーテッドアニーリングの技法がある。それについては12.1節を参照されたい。

話をはっきりさせるために、0と1の二値をとりうる素子が結合したネットワークを考えよう。このネットワークに対し、エネルギー関数

$$E(\{S\}) = \sum_{\langle i,j \rangle} J_{ij} S_i S_j$$

が定義されているとする。 S_i は*i*番目の素子の値であり、和は結合しているすべての素子対についてとる。ここでは、素子間の結合定数 J_{ij} があらかじめすべて決められており、動的に変化しないものとしておく。この系では、すべての素子についてそれぞれ0と1のどちらの値をとるかを決めてやれば、系の微視的配位がひとつ指定されたことになる。 $E(\{S\})$ という表記は、各微視的配位に対してエネルギー関数の値が一意に決まることを表している。

熱平衡状態とは各微視的配位がそのエネルギー値に応じて

$$P_{\text{eq}}(\{S\}; T) = \frac{1}{Z} e^{-E(\{S\})/T}$$

という重みで含まれる(ボルツマン分布する)状態をいう。 T は温度と呼ばれ、外部から指定するパラメータである。明らかに、高温極限ではすべての配位が等しい重みで実現し、低温極限では最低エネルギー状態のみが実現する。規格化定数

$$Z = \sum_{\text{全配位}} e^{-E(\{S\})/T}$$

は分配関数と呼ばれる。重みがエネルギー値だけで決まることが重要で、これにより微視的には異なっても同じ重みをもつ配位が多数ありうる。同じエネルギー値をもつ微視的配位の数 $\Omega(E)$ と書けば、エネルギー E の実現確率は

$$P_{\text{eq}}(E) = \frac{\Omega(E) e^{-E/T}}{Z}$$

となる。

さらに熱平均の概念を導入する。各微視的配位ごとに値が一意に決まるような物理量 $Q(\{S\})$ の熱平均は

$$\langle Q \rangle = \sum_{\text{全配位}} P_{\text{eq}}(\{S\}; T) Q(\{S\})$$

で与えられる。この熱平均を計算するのが、モンテカルロ法の目的である。

素子数が少ない場合には全微視的配位を枚挙することによって $P_{\text{eq}}(\{S\}; T)$ を完全に求めることができ、従って熱平均も容易に計算できる。しかし、可能な微視的配位の数 $\Omega(E)$ が素子数 N に対して 2^N であることから明らかなように、素子数とともにこの方法は急激に難しくなる。モンテカルロ法では、この困難を回避するために、全配位を数え上げることをあきらめ、代表的な微視的配位のみを確率的に生成して、それらによる平均をもって熱平均の近似値とする。このとき代表的な配位をどのように選び出すかが問題となる。微視的配位の重みがエネルギーの指数関数で与えられることから、一様なランダムサンプリングは用をな

さない。代わりに、各微視的配位を平衡分布 $P_{\text{eq}}(\{S\}; T)$ に比例する確率で抽出しようというのがメトロポリス法である。もちろん、平衡分布を求めるのが目的なのだから、 $P_{\text{eq}}(\{S\}; T)$ があらかじめわかっているわけではないが、実は確率過程を用いることによりこれが実現されるのである。

そこで、定常分布が熱平衡分布を与えるような時系列を離散時間のマルコフ連鎖によって作ることを考えよう。つまり、時系列に関する時間平均が熱平均と一致して

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M Q_m = \langle Q \rangle$$

となるようにするようになるわけである。 Q_m は m ステップ目での Q の値を表す。マルコフ連鎖であるから、ある時刻での系の状態は直前の時刻での状態のみで決まり、それ以前の歴史にはよらない。従って、系の取りうるすべての微視的配位に対し、次の時刻に他の配位に遷移する確率をすべて与えればマルコフ連鎖が定義される。配位 μ から配位 ν に遷移する確率を $w_{\nu\mu}$ とすると、これは確率保存から

$$0 \leq w_{\nu\mu} \leq 1$$

$$\sum_{\nu} w_{\nu\mu} = 1$$

を満たすように決められる。遷移確率 $w_{\nu\mu}$ を要素とする行列 W がマルコフ行列なので、その最大固有値は 1 で対応する固有ベクトルは非縮退である。したがってこの確率過程は適当な初期分布 $P_0(\{S\})$ から出発して、長時間極限で定常分布 $P_{\infty}(\{S\})$ を与える。

ここで、この定常分布が初期分布の取り方によらず一意に決まることが保証されなくてはならない。そのための条件は、配位空間がいくつかの副空間に分離しないこと、すなわち、系の任意の微視的配位から他の全ての配位へ有限回のステップで遷移する道筋が存在することである。これはエルゴード性の条件と呼ばれる。今のような素子系の場合には、1 素子の値だけが変化するような遷移を考えることにすれば、エルゴード性は自明に成り立たせることができる。もちろん、一般の場合にエルゴード性がなりたつかどうかは自明ではない。

さらに、この定常分布 $P_{\infty}(\{S\})$ が熱平衡分布 $P_{\text{eq}}(\{S\}; T)$ に一致するためには、 W が詳細釣り合いの条件

$$w_{\nu\mu} P_{\text{eq}}(\mu; T) = w_{\mu\nu} P_{\text{eq}}(\nu; T)$$

を満たせばよい。あるいは、書き直して

$$\frac{w_{\nu\mu}}{w_{\mu\nu}} = \frac{P_{\text{eq}}(\nu; T)}{P_{\text{eq}}(\mu; T)} = e^{-(E(\nu) - E(\mu))/T}$$

である。ところが、これは $w_{\nu\mu}$ と $w_{\mu\nu}$ の比だけに関する制限なので、これを満たす W は一通りではない。この条件を満たす限り、どんな W を使っても同じ熱平衡分布が実現する。

具体的な W の形については以下で述べることにして、とりあえず、ここまでで熱平衡分布を与える確率過程が定義されたとしよう。この確率過程について計算機シミュレーションを行い、熱平衡分布を実際に作り出すのが次の仕事である。実際のシミュレーションでは、適当に選んだ初期配位から遷移確率に従って次々と配位間を遷移させることにより、時系列を発生させる。十分に長いシミュレーションの後には、初期状態によらず分布関数が平衡分布に達するはずなので、その後の時系列は平衡分布関数からのサンプリングになる。これにより、得られた時系列での Q の平均値は期待値の意味で熱平均に一致するわけである。

12.5.2 例

ここでは、代表的なふたつの遷移確率のとりかたについて、二値素子系での具体的な手順を述べる。なお、簡単のために一個の素子の値だけが違うような配位間の遷移だけを許すことにする。この場合、上述の通り、エルゴード性は自明になりたっている。

状態変更の対象となる素子の選びかたには任意性があり、ランダムに選んでも規則的に選んでもかまわない。使用する乱数が少なくすむという点では規則的に選んだほうが有利である。もっとも、1ステップの実行に要する時間が短いからといって、必要な精度の結果を短い時間で得られるとは限らない。本当の意味で有利かどうかは場合によると思われる。

(狭義の)メトロポリス法 この方法では、遷移先の状態 ν をひとつ選び、現在の状態 μ とのエネルギー差から遷移確率を次のようにとる。

$$w_{\nu\mu} = \begin{cases} 1 & (E_\nu < E_\mu) \\ e^{-(E(\nu)-E(\mu))/T} & (\text{それ以外}) \end{cases}$$

遷移確率がエネルギー差だけで書けているので、いちいち全エネルギーを計算する必要はない。

二値素子系での手順はたとえば以下のようになる。

1. 状態変更の試行を行う対象として一個の素子を選び、状態変更前後のエネルギー差 ΔE を計算する。二値素子なので、現在の値が0なら変更後の値は1、現在の値が1なら変更後は0と一意に決まる。なお、多値素子の場合は変更後の値を乱数によって選ぶ。
2. $[0, 1]$ の一様乱数 r を発生させ、遷移確率 $w_{\nu\mu}(\Delta E)$ と比較する。 $r < w_{\nu\mu}$ なら試行が成功したものととして素子の値を変更し、 $r > w_{\nu\mu}$ なら試行が失敗したものととして変更しない。
3. 上の手続きを繰り返す。

熱浴法 この方法では、試行対象となる素子を選んだら、その素子のとりうるすべての状態(現在の状態も含めて)への遷移確率を以下のように作る。

$$w_\nu = \frac{e^{-E(\nu)/T}}{\sum_\gamma e^{-E(\gamma)/T}}$$

ただし、分母の和は素子のとりうる全状態についてとる。この方法では確率が現在の素子の値によらず、変更後の状態だけで決まる。二値素子の場合には、素子のとりうる値が0と1だけなので

$$\begin{cases} w_0 & = \{1 + e^{-(E(1)-E(0))/T}\}^{-1} \\ w_1 & = 1 - w_0 \end{cases}$$

となり、メトロポリス法同様ふたつの状態のエネルギー差だけで書ける。この遷移確率はボルツマンマシンの遷移確率にほかならない。

熱浴法による二値素子系のシミュレーションの手順は例えば以下のようになる。

1. 状態変更の試行を行う対象として一個の素子を選び、それが0のときと1のときのエネルギー差 ΔE を計算する。
2. $[0, 1]$ の一様乱数 r を発生させる。 $r < w_0(\Delta E)$ なら素子の値を0とし、 $r > w_0(\Delta E)$ なら1とする。
3. 上の手続きを繰り返す。

12.5.3 一般的な注意

ここで、実際にモンテカルロシミュレーションを行うに際の注意をいくつか挙げておく。

サンプリングの頻度 確率過程が定義できたとして、その中からどういう頻度でサンプルをとればいいのかを考えてみよう。前節の例では1ステップにつき最大1素子しか状態が変わらないので、毎ステップのデー

タをとっても統計精度をあげるという観点からはほとんど意味がない。普通は、全素子について一度ずつ（ランダムに素子を選んでいる場合は平均一度）の試行を行ったときを1モンテカルロステップとよび、時間をこの単位ではかることが多い。これは1モンテカルロステップ以下の間隔でサンプリングしても有効ではないと思われるからである。何モンテカルロステップごとにサンプルをとるかは、以下で議論する平衡緩和の緩和時間を考慮して、1モンテカルロステップの実行に要する時間と1サンプルについて物理量の計算をするのに要する時間との兼ね合いで決めればよい。

初期緩和 適切な初期配位から出発するので、明らかにシミュレーションの最初のほうは平衡分布からのサンプリングになっていない。一般に物理量はステップ数に対して（少なくとも漸近的には）指数関数的に緩和するので、この初期緩和部分のデータは熱平均の計算から除いておかななくてはならない。つまり、平衡分布に達したことを確認してから、熱平均を計算するためのサンプルをとるという手続きが必要となる。

平衡緩和 平衡分布に達したあとであっても、確率過程を用いているために、相続いて得られるサンプルのあいだには強い相関がある。一般にふたつの時刻で得られた Q の値の相関は漸近的に

$$\langle Q_m Q_n \rangle \sim e^{-|m-n|/\tau}$$

とふるまう。 τ は緩和時間とよばれる。したがって、シミュレーションによって M サンプルを得たとしても、それをすべて統計的に独立なサンプルとみなして統計処理をすることはできない。十分に長いシミュレーションを行ったとすると、実効的な独立サンプルの数は

$$M_{\text{eff}} = \frac{M}{2\tau}$$

で与えられる。統計的な精度は独立サンプルの数で決まるので、緩和時間を見積もった上で必要な計算ステップ数を決める手続きを踏むべきだろう。特に物理量の分散や共分散を求めようとする、独立サンプル数によるバイアスを考慮する必要が出てくるので、独立サンプル数の正確な見積もりが要求される。なお、緩和時間は計算する物理量によっても異なる。

初期緩和にしる熱平衡での緩和にしる、問題によっては最終的な指数緩和に漸近するまでに長時間を要することがある。そのような遅い緩和の原因は問題によりけりで、それを回避する万能の処方は今のところない。問題によっては、緩和の速い等価な問題に変換するなどの工夫ができることもある。また、緩和の遅さが配位間のエネルギー障壁に起因する場合には、あとの節でとりあげるマルチカノニカル法などの方法が効果的かもしれない。

採択率 試行の採択率（状態変更の成功率）は、計算効率のひとつの目安を与える。いかに1ステップあたりの実行時間が短くても、採択率が低ければ実効的に得られる独立サンプルの数は少なくなってしまう。種類の異なる遷移をいくつか組み合わせている場合などは、特に注意が必要である。極端に採択率の低い遷移が含まれる場合は、その遷移が本当に必要なものかを考え、問題の性質上不可欠であるならば、その遷移が充分多数回行われるまでシミュレーションを続けるか、さもなければアルゴリズムを見直すべきだろう。

前節では一素子だけが値を変えるような遷移のみを扱った。もちろん、もっと多くの素子の値を一度に変えるような遷移を考えてもかまわない。ただし、やみくもにそのような遷移をとりいれても、普通は試行の採択率が急激に下がるので必ずしも得にはならない。問題の特性に応じたうまい遷移を工夫して採択率をあげる試みもなされている。

12.5.4 乱数

モンテカルロ法を使うにあたっては乱数の選択が重要である。乱数には大きくわけて、熱揺らぎなどの物理現象を利用する自然乱数とアルゴリズムによって生成される擬似乱数がある。自然乱数発生器を手元に置いているケースはまれだと思うので、ここでは擬似乱数について特に使用上の注意点を述べておく。

真の乱数であれば、あらゆる意味で無相関でなくてはならないし、当然同じ数列の再出現（回帰）などが起きてはならない。しかしながら、アルゴリズムを用いてそのような意味で真の乱数を生成するのは原理的に不可能である。したがって、なるべくよい擬似乱数を選択することが問題となる。とはいっても、擬似乱数の「よさ」を一般的に定義することはできない。実際、ある目的のためには充分によい乱数として働くが、別の目的では思いがけない悪さをするような例もまま見受けられる。これについての一般的な処方は今のところなく、扱う問題ごとにいろいろと試すしかなさそうである。以下によく用いられる乱数の例を挙げる。いずれの方法を使うにしても、所詮はアルゴリズムで生成した擬似乱数列であり、万能なものではありえないことを常に念頭に置くべきである。

乗算合同法 これは

$$r_{n+1} = ar_n \pmod{b}$$

という漸化式から整数乱数列 $\{r_i\}$ を生成するものである。 a と r_1 は奇数にとる。 a と b の値の選びかたは乱数の性質に大きく影響し、「よい」とされる値について多くの提案がなされている。

アルゴリズムから明らかとおり、この方法で生成される乱数列は $b/2$ 以下の周期を持つ。最近のシミュレーションでは 2^{30} 程度の乱数は容易に使い尽くしてしまうので、語長が 32 ビット程度のものは乱数として不適当な場合が多い。整数の語長が充分長い場合、例えば 64 ビット以上なら乱数の回帰を問題にする必要は事実状ないだろう。

周期より深刻な問題として、相続いて得られる値の組を多次元空間の座標値とみなして点を打つと、点はランダムには打たれずいくつかの面に乗ることが知られている。したがって、少なくとも座標値を得る目的でこの方法を用いるべきではないし、問題によってはこの相関が思わぬ悪さをする可能性がある。また、この方法では下位ビットにいくほど周期が短くなるという点にも注意が必要である。したがって、上位ビットをマスクして下位ビットだけを用いるような使い方には向いていない。得られた乱数列をさらにかきまぜて使うシャッフル法は比較的たちのよい乱数列を生成するようである。

M 系列 合同法に代わって最近よく使われる擬似乱数列として、いわゆる M 系列がある。これは n 個のビット列から

$$r_{n+1} = \text{XOR}(r_{n-p}, r_{n-q})$$

によって 1 ビット擬似乱数列、すなわち 0 と 1 のランダムな列が生成されることを用いる。その際、 $x^p + x^q + 1$ が GF(2) 上の原始三項式で、かつ $2^p - 1$ が素数となるように $p, q (p > q)$ を選ぶことにより、最大周期 $2^p - 1$ が実現される。このような p, q の組はすでに表として与えられている。こうして作られたビット列を適当な長さに区切れればランダムな整数（の 2 進表現）列が得られる。

実際には、1 ビット乱数を順に発生させる代わりに、必要な語長分だけビット列を並べておいて排他的論理和を並列にとる方法がよく用いられる。この場合、新しい乱数をひとつ得るためには排他的論理和を一度実行するだけである。

この方法では、 p として大きな値を選んでおけば、たちのよい乱数列が得られると考えられている。また、上位ビットから下位ビットまでが 1 ビット乱数として同じ性質をもつという特徴もある。ただし、最大周期列を実現するには「乱数の種」に相当する初期ビット列の作り方に注意が必要である。

12.5.5 新しい技法

低エネルギーの配位が多数あり、その間が高いエネルギー障壁で隔てられているような場合には、単純なモンテカルロ法ではどれかひとつの低エネルギー配位付近に凍結して抜けられなくなってしまい、うまく低温での熱平衡状態にたどりつけないことがある。この問題を解決するためにいくつかの方法が提案されている。この節では最近発展してきたいくつかの新しい技法について述べる。

ヒストグラム法 メトロポリス法では、各微視的配位がボルツマン分布にしたがう確率で得られる。その際、温度は外部から与える固定されたパラメータであった。しかしながら、各配位の出現頻度が得られるのであるから、サンプルを得たのちにボルツマン分布の重率を変えることにより、外部から与えたのとは別の温度での熱平均を計算することも原理的にはできるはずである。それにはモンテカルロシミュレーションによって得られた各配位の実現確率 $P(\{S\}; T)$ に対して

$$P(\{S\}; T') = \Omega(E) e^{-E(\{S\})/T'} \propto P(\{S\}; T) e^{E(\{S\})(1/T - 1/T')}$$

の関係をいれればよい。エネルギー分布にだけ興味があるなら、シミュレーションでエネルギーヒストグラム $P(E; T)$ を作り、

$$P(E; T') \propto P(E; T) e^{(E/T - E/T')}$$

とする。

この方法は、新しい温度 T' が T に充分近ければうまく働く。一方サンプル数が有限であるために、 T' が T から離れるにつれて、結果にはバイアスがかかるようになるので注意が必要である。どの程度の温度範囲まで有効かは、エネルギー分布の広がりから見積もることができる。一般に、素子数 N に対し

$$\sqrt{\langle E^2 \rangle - \langle E \rangle^2} \sim O(\sqrt{N})$$

であるから、有効な温度領域の幅は素子数とともに $1/\sqrt{N}$ で狭くなる。

また、バイアスを避けるために、いくつかの温度でヒストグラムを作り、温度に関する内挿にヒストグラム法を用いることも提案されている。

マルチカノニカル法 メトロポリス法の原理に立ち戻ってみると、ボルツマン分布が実現したのは、結局詳細釣り合いの式に平衡分布としてボルツマン分布の形をいれたためである。とすれば、ここに別の形の分布をいれれば別の平衡分布が実現できるはずである。実際、エネルギーの勝手な関数 $f(E)$ を使って、平衡分布として

$$P_{\text{eq}}(\{S\}) \propto e^{-f(E(\{S\}))}$$

を要求してもよく、こうして得られたエネルギー分布 $P(E)$ から、上述のヒストグラム法を応用して適当な温度でのボルツマン分布が構成できる。マルチカノニカル法と呼ばれる方法では、エネルギーヒストグラムがなるべく平らになるように $f(E)$ を決める。すなわち

$$P_{\text{eq}}(E) \propto \frac{1}{\Omega(E)}$$

となることを要請する。本来ならエネルギーヒストグラムの形を知るには、遷移確率を決めた上でシミュレーションを行わなければならないはずなので、これは一見実現できそうにない。しかし、実用上は上式が厳密になりたなくてもいいので、短いシミュレーションを繰り返してヒストグラムを順次改良するように $f(E)$ の形自体を学習させることにより、これを近似的に実現できる。

$f(E)$ の形としては、 E に対して区分線形な折れ線を用いる場合が多い。こうすると、形式的にはエネルギー範囲ごとに違う温度で計算を行っていることに対応し、直感的にもわかりやすい。これはまた、 $\log \Omega(E)$

を折れ線で近似したことにもなっている。ヒストグラム法での見積もりと同様に、ひとつの温度で賄える温度領域は素子数とともに $1/\sqrt{N}$ で狭くなるため、それにしがって $f(E)$ を記述する折れ線を細かくしていかななくてはならない。

この方法ではシミュレーションがエネルギー空間でのランダムウォークを与えるので、配位間に高いエネルギー障壁がある場合に有効と考えられている。

交換法 マルチカノニカル法がひとつのシステムをエネルギー値に応じて違う温度でシミュレーションしようとするのに対し、同様に幾通りもの温度を用意したうえ更に温度の数だけシステムも用意する方法が考えられる。もちろん、システムごとに温度を決めてしまえば単に普通のモンテカルロシミュレーションをいくつも実行するだけのことだが、一定の間隔でシステムそのものを交換してしまうプロセスを取り入れることにより、マルチカノニカル法と同様に配位間のエネルギー障壁をうまく乗り越えようというのが、この交換法である。

簡単のため、ふたつの同一なシステムを異なる温度 T_1 と T_2 で同時に計算することを考えよう。システムがともに熱平衡状態にあるとすると、エネルギーがそれぞれ E_1 と E_2 である確率は

$$P(E_1, E_2; T_1, T_2) = P(E_1; T_1)P(E_2; T_2) = e^{-E_1/T_1 - E_2/T_2}$$

で与えられる。このふたつのシステムをとり替える（あるいは温度をとり替えると思ってもよい）ような遷移を導入してもふたつのシステムが同時にそれぞれの温度で熱平衡にあるためには、システムのとりに替える遷移確率について次の詳細平衡がなりたてばよい。

$$\frac{w(1, 2 \rightarrow 2, 1)}{w(2, 1 \rightarrow 1, 2)} = e^{-(1/T_1 - 1/T_2)(E_2 - E_1)}$$

したがって、それぞれのシステムごとに通常のモンテカルロシミュレーションを行いながら、適当な頻度でメトロポリス法あるいは熱浴法によってこのシステムのとりに替えを挟みこめば、ふたつの温度での平衡分布が同時に得られる。

この考え方は同時に扱う温度の数を増やしても変わらない。十分に高温から低温までの温度を用意すれば、エネルギー障壁を超えていくつもの低エネルギー状態を実現できると考えられる。なお、温度はシステムの入れ替えが充分頻繁に行われるように設定する必要がある。そのためには、温度の間隔をマルチカノニカル法と同様に $1/\sqrt{N}$ で細かくしていかななくてはならない。マルチカノニカル法と違い、温度の上限・下限を一定に保つとすれば、これは同時に扱うシステムの数を \sqrt{N} で増やすことを意味する。

12.5.6 参考文献

標準的な教科書

Monte-Carlo Simulation in Statistical Physics — An Introduction, 2nd. edition : K. Binder and D. W. Heermann (Springer, 1992)

独立サンプル数の推定

M. Kikuchi, N. Ito and Y. Okabe: in *Computer Simulations in Condensed Matter Physics VII* ed. D.P. Landau, K.K. Mon and H.B. Schüttler (Springer, 1994) 44.

乱数

乱数：伏見正則（東大出版会、1989）

The Art of Computer Programming, Vol. 2, 2nd. edition: D.E. Knuth (Addison-Wesley, 1981) (準数値算法：サイエンス社、1981)

Numerical Recipes, 2nd. edition: W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery
(Cambridge, 1992)

ヒストグラム法

A. M. Ferrenberg and R. H. Swendsen : Phys. Rev. Lett. **61** (1988) 2635.; Phys. Rev. Lett. **63**
(1988) 1195.

マルチカノニカル法

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