イジングモデルのモンテカルロ法

須賀勇貴

最終更新日：2023 年 11 月 8 日

ここではイジングモデルに対するメトロポリス法と熱浴法についてまとめる．主に富谷先生の「これならわかる機械学習入門」の 10 章を参考にして作成した．

モンテカルロ法とは乱数を使って積分や和を行う手法である．ここでは，期待値の和を乱数を用いて調べる．あるスピン配位 C の実現確率を

$$P(C) = \frac{1}{Z} e^{-\beta E[C]} \tag{1}$$

として確率的にスピン配位 C を生成して，生成したスピン配位 C を使った平均を用いて期待値を推定する手法をとる．

しかし，スピン配位 C の実現確率 (1) は直接計算できない．なぜなら，分母の分配関数 Z を計算するために，結局全ての配位の和を計算する必要があるからである．そこで，分配関数 Z の計算を避けてスピン配位を生成できつ手法であるマルコフ連鎖モンテカルロ法 (Markov chain Monte-Carlo) を以下で説明する．

1　マルコフ連鎖モンテカルロ法

今，調べたい期待値は一般的に

$$\mathbb{E}[f(C)] = \sum_C f(C) P(C) \tag{2}$$

と書ける．ここで $f(C)$ は配位の関数で，例えば磁化率などである．また，$\sum_C$ は可能なスピン配位についての足し上げを表す．格子点の数が K 個であるとき，イジングモデルなら 2^K 個の足し算になる．これを

$$\mathcal{C} = \{C_1, C_2, C_3, \cdots\} \tag{3}$$

という 2^K よりもずっと短い列の平均に置き換え，

$$\mathbb{E}[f(C)] \approx \frac{1}{|\mathcal{C}|} \sum_{C_k \in \mathcal{C}} f(C_k) \tag{4}$$

を代わりに評価したい．ここで $|\mathcal{C}|$ は配位の数を表す．当然，適当に作った列では正しい期待値を与えないが，以下を満たす $P(C_a|C_b)$ に従って C を生成すれば正しい期待値を与えることが知られている．

$$P(C_a|C_b)P(C_b) = P(C_b|C_a)P(C_a) \tag{5}$$

この証明は大数の法則による．ここで条件付き確率 $P(C_a|C_b)$ は遷移確率とよばれ，この式は詳細釣り合い (detailed balance) の式と呼ばれている．乱数を使ったアルゴリズムをモンテカルロ法と呼ぶが，上のような確率的に配位の列を作るアルゴリズムをマルコフ連鎖モンテカルロ法という．

イジングモデルに対するメトロポリス法

マルコフ連鎖モンテカルロ法の中で最もシンプルな手法であるメトロポリス法について見ていく．ある配位 C_b と C_a を比較したときに C_a のエネルギーの方が低い場合, つまり $E(C_b) > E(C_a)$ を考える．この場合は C_a の方がエネルギー的に安定なので，C_b から C_a への遷移確率を

$$P(C_a|C_b) = 1 \tag{6}$$

ととることにする．このとき，詳細釣り合いの式は

$$P(C_b) = P(C_b|C_a)P(C_a) \tag{7}$$

となる．これより，C_a から C_b の遷移確率が決まってしまい，

$$P(C_b|C_a) = \frac{P(C_b)}{P(C_a)} = \frac{e^{-\beta E(C_b)}/Z}{e^{-\beta E(C_a)}/Z} = e^{-\beta\left(E(C_b)-E(C_a)\right)} \tag{8}$$

となる．

したがって，C_a から C_b の遷移確率として，

$$P(C_b|C_a) = \begin{cases} \exp\Big[-\beta\Big(E(C_b)-E(C_a)\Big)\Big], & E(C_b) > E(C_a) \\ 1, & E(C_b) < E(C_a) \end{cases} \tag{9}$$

ととると詳細釣り合いの式を満たすことになる．このアルゴリズムをメトロポリス法と呼ぶ．

メトロポリス法をイジングモデルに適用する場合の手順をまとめる．

1. 適当な配位 C_1 を用意する．
2. 以下を $i = 1, 2, 3, \cdots, N_{\mathrm{conf}}$ まで繰り返す，
   (a) C_i のあるサイトのスピンを反転し，その配位を C_i' とおく．
   (b) エネルギー $E(C_i)$ と $E(C_i')$ を計算する．
   (c) $E(C_i) > E(C_i')$ なら $C_{i+1} = C_i'$ とし，$i \to i+1$ とおいて (a) へ戻る．(受理 (accept))
   (d) 乱数 $r \in [0,1]$ を用意し，$\Delta E = E(C_i') - E(C_i)$ に対して，$r < \exp(-\beta\Delta E)$ ならば $C_{i+1} = C_i'$ とし，$i \to i+1$ とおいて (a) へ戻る．(受理 (accept))
   (e) $C_{i+1} = C_i$ とし，$i \to i+1$ とおいて (a) へ戻る．(棄却 (reject))

ここで，「棄却」は提案された配位を使わずに前の配位をそのまま次の配位として受け入れるという意味であることに注意してほしい．

上記の手順に従うと，配位の列

$$\mathcal{C} = \{C_1, C_2, C_3, \cdots, C_{N_{\mathrm{conf}}}\} \tag{10}$$

が得られる．しかし最初の方の配位は適当に選んだ初期配位の影響を受けているため，期待値の計算に含めてはいけない．例えば $N_{\mathrm{disc}}(1 < N_{\mathrm{disc}} << N_{\mathrm{conf}})$ 番目 (disc は discard の略) まで除いたとすると，

$$\mathbb{E}[f(C)] \approx \frac{1}{N_{\mathrm{disc}} - N_{\mathrm{conf}}} \sum_{k=N_{\mathrm{disc}}}^{N_{\mathrm{conf}}} f(C_k) \tag{11}$$

のように期待値を計算できる．N_{disc} は配位ごとの物理量を観察して初期配位の影響がなくなったあたりの番号をとればよい．

イジングモデルの熱浴法

次に熱浴法 (heatbath algorothm) をみる．このアルゴリズムは機械学習の文脈ではギブスサンプリング (Gibbs sampling) と呼ばれる．熱浴法は，ある 1 自由度を着目し，その周りの自由度を熱浴とみなして更新するマルコフ連鎖モンテカルロ法の一種である．

熱浴法は条件付き確率

$$P(A|B) = \frac{P(A,B)}{P(B)} \tag{12}$$

から導出される．イジングモデルのあるサイト j に着目し，そのサイト以外を熱浴をみなしたときにサイト j がどうなるかを条件付き確率に従って決めるのである．

例えば,

$$A = A_j\text{: サイト } j \text{ にあるスピンの状態} \tag{13}$$

$$B\text{: サイト } j \text{ 以外のスピンの状態} \tag{14}$$

とする．ある配位を与える確率を

$$P(A_j, B) = \frac{1}{Z} \exp\left[-\beta J s_j \sum_{k\in\mathcal{N}(j)} s_k + (s_j\text{を含まない項})\right] \tag{15}$$

と書く．ここで，サイトをまとめて $\mathcal{N}(j)$ と書いた．

A_j の状態 (サイト j のスピン状態) を足し上げて周辺化した確率 $P(B)$ は，$P(B) = \sum_{A_j} P(A_j, B)$ を使うと A_j は $s_j = \{+1, -1\}$ をとるので，全て足し上げて

$$P(B) = \frac{1}{Z} \sum_{s_j=\pm 1} \exp\left[-\beta J s_j \sum_{k\in\mathcal{N}(j)} s_k + (s_j\text{を含まない項})\right] \tag{16}$$

とわかる．ここから条件付き確率 $P(A_j|B)$ は，定義 $P(A_j|B) = P(A_j, B)/P(B)$ から

$$
\begin{aligned}
P(A_j|B) &= \frac{1}{Z}\exp\left[-\beta J s_j \sum_{k\in\mathcal{N}(j)} s_k + (s_j\text{を含まない項})\right] \Bigg/ \\
&\quad \frac{1}{Z}\sum_{s_j=\pm1}\exp\left[-\beta J s_j \sum_{k\in\mathcal{N}(j)} s_k + (s_j\text{を含まない項})\right] \\
&= \frac{\exp\left[-\beta J s_j \sum_{k\in\mathcal{N}(j)} s_k\right]}{\sum_{s_j=\pm1}\exp\left[-\beta J s_j \sum_{k\in\mathcal{N}(j)} s_k\right]} \\
&= \frac{\exp\left[-\beta J s_j \sum_{k\in\mathcal{N}(j)} s_k\right]}{\exp\left[-\beta J \sum_{k\in\mathcal{N}(j)} s_k\right] + \exp\left[+\beta J \sum_{k\in\mathcal{N}(j)} s_k\right]}
\end{aligned}
\tag{17}
$$

となる．最後の式を見ると，サイト j の周りのスピン配位だけでサイト j の状態が決まるということがわかる．また，ここでもメトロポリス法のときと同様に分配関数 Z の計算が必要なくなっている．

例えば，サイト j のスピンが $+1$ である場合は

$$
P(s_j = +1) = \frac{\exp\left[\beta J \sum_{k\in\mathcal{N}(j)} s_k\right]}{\exp\left[\beta J \sum_{k\in\mathcal{N}(j)} s_k\right] + \exp\left[-\beta J \sum_{k\in\mathcal{N}(j)} s_k\right]} \tag{18}
$$

となる．なのでこの確率に従って次の配位におけるサイト j のスピンを決定 (更新) すればよい．この確率の評価はメトロポリス法の乱数を使っている個所と同じように乱数を用いて評価すればよい．また $P(s_j = -1|B) = 1 - P(s_j = +1|B)$ である．そして空間全部のスピンについて繰り返せば次の配位が得られる．