

量子モンテカルロ法について

筒井翔一郎

2022 年 1 月 13 日

目次

1	モンテカルロ数値積分	1
1.1	具体例 1	2
2	ハイブリッドモンテカルロ法	2
3	再重み付け法と符号問題	3
4	符号問題を回避する方法	4
4.1	テイラー展開法	4
4.2	解析接続	4
4.3	複素ランジュバン法	5
4.4	Lefschetz thimble	7
5	Summary	8

1 モンテカルロ数値積分

一般に、多自由度の積分を区分求積によって数値的に計算することは困難である。このような場合に利用されるのがモンテカルロ法である。いま、対象の積分が

$$I = \int dx_1 dx_2 \cdots dx_n O(x_1, x_2, \cdots, x_n) P(x_1, x_2, \cdots, x_n) \quad (1)$$

という形をしていて、 $0 \leq P \leq 1$ かつ $\int dx_1 dx_2 \cdots dx_n P(x_1, x_2, \cdots, x_n) = 1$ を満たしているとする。このとき、 P は確率密度変数とみなすことができるから、積分 I は P に従う確率変数 X_1, X_2, \cdots, X_n に対する $O(X_1, X_2, \cdots, X_n)$ の期待値であると考えてもよい。従って、積分 I は

$$I = \frac{1}{N} \sum_{j=1}^N O(X_1^j, X_2^j, \cdots, X_n^j) + \mathcal{O}(1/\sqrt{N}) \quad (2)$$

と評価することができる。ここで、 N はサンプル数、 $\{X_1^j, X_2^j, \cdots, X_n^j\}$ は P に従う確率変数で、 $X_k^j, X_k^{j'} (j \neq j')$ は独立である。 P が振る舞いの良い確率密度関数ならば、大数の法則から O の誤差はガウス分布から見積もることができ、そのオーダーは $\mathcal{O}(1/\sqrt{N})$ となる。モンテカルロ法の優れた点は以下である。

1. P がある領域に鋭いピークをもつような関数の場合、 I を評価するにはその領域近傍を考えれば十分であり、定義域全体をくまなく考慮する必要はない。その点モンテカルロ法は、積分に寄与する領域のみを重点的にサンプリングするようになっているため、効率よく I を評価できる。
2. I の誤差は統計誤差であるから、コントロールしやすい。

1.1 具体例 1

積分

$$I = \frac{\int_{-\infty}^{\infty} dx x^2 e^{-x^2}}{\int_{-\infty}^{\infty} dx e^{-x^2}} \quad (3)$$

にモンテカルロ法を適用するには、 $O(x) = x^2$, $P(x) = e^{-x^2} / \int_{-\infty}^{\infty} dx e^{-x^2}$ とすれば良い。

2 ハイブリッドモンテカルロ法

モンテカルロ法を実行するためには、確率密度関数 P に従う確率変数を発生させる必要がある。これを実現するアルゴリズムのひとつが、ハイブリッドモンテカルロ (HMC) 法である。以下では P として、

$$P(x_1, x_2, \dots, x_n) \propto e^{-S(x_1, x_2, \dots, x_n)} \quad (4)$$

という関数形を仮定する。ちなみに、 $P \propto f e^{-S}$, ($f > 0$) というタイプの確率密度関数も、 $e^{-S+\log f}$ と書き直して $S - \log f$ を新たに S と読み替えることで上の形に帰着するので、この仮定の範囲内である。

次に、以下のような連立微分方程式

$$\frac{dx_i(t)}{dt} = \frac{\partial H}{\partial p_i(t)} = p_i(t) \quad (5)$$

$$\frac{dp_i(t)}{dt} = -\frac{\partial H}{\partial x_i(t)} = -\frac{\partial S}{\partial x_i(t)} \quad (6)$$

を考える。。ここで H は

$$H(p_i, x_i) = \frac{1}{2} \sum_{i=1}^n p_i^2 + S(x_1, x_2, \dots, x_n) \quad (7)$$

で与えられるものとする。 H をハミルトニアンと呼び、上の微分方程式をハミルトン系^{*1}、あるいは運動方程式と呼ぶ。この方程式を、適当な $\{x_i(0)\}$ とランダムに発生させた $\{p_i(0)\}$ を初期条件として $t=0$ から $t=1$ まで解き、得られた $\{x_i(1)\}$ をサンプリングするという手順を繰り返すことで、平衡分布が P であるようなマルコフ連鎖を作ることができる。ただし、これを計算機上で実現するためには、運動方程式を適切なスキームで離散化して解き、なおかつメトロポリステストと呼ばれる操作を行って、数値誤差の影響を補正する必要がある。このことを念頭においた上で具体的なアルゴリズムを書き下すと、以下のようになる^{*2}。

1. $x_i(0)$ を適当にセットする。

^{*1} ハミルトン系を用いるという点から、この手法のことをハミルトニアンモンテカルロ (Hamiltonian Monte Carlo) 法と呼ぶ儀もある。

^{*2} 手前味噌ですが、[このページ](#)でサンプルコードを公開しています。

2. $p_i(0)$ を標準正規分布に従って発生させる。
3. 運動方程式を $t = 0$ から $t = 1$ まで解く。時間刻み幅を $\Delta t = 1/N_t$ とおく。 $x_i(l\Delta t) = x_i^l$, $p_i(l\Delta t) = p_i^l$ と略記する。離散化された運動方程式は以下で与えられる。

$$p_i^{1/2} = p_i^0 - \frac{1}{2} \frac{\partial S}{\partial x_i^0} \Delta t \quad (8)$$

$$x_i^{k+1} = x_i^k + p_i^{k+1/2} \Delta t \quad (9)$$

$$p_i^{k+3/2} = p_i^{k+1/2} - \frac{\partial S}{\partial x_i^{k+1}} \Delta t \quad (10)$$

$$x_i^{N_t} = x_i^{N_t-1} + p_i^{N_t-1/2} \Delta t \quad (11)$$

$$p_i^{N_t} = p_i^{N_t-1/2} - \frac{1}{2} \frac{\partial S}{\partial x_i^{N_t}} \Delta t \quad (12)$$

ここで、 $k = 0, 1, \dots, N_t - 2$.

4. r を区間 $[0, 1]$ の一様分布に従って発生させた確率変数とする。 r が不等式

$$r < \min\{1, e^{-\Delta H}\}, \quad \Delta H = H(p_i^{N_t}, x_i^{N_t}) - H(p_i^0, x_i^0) \quad (13)$$

を満たすとき、 $x_i^{N_t}$ をサンプリングし、これを新たに x_i^0 とする。満たさなかった場合はなにもしない。

5. 上のステップを繰り返す。

3 再重み付け法と符号問題

モンテカルロ法の重要な仮定は、 P が正定値であることだった。実は P が正定値でない場合でも、形式的にはモンテカルロ法はいつでも適用可能である。

いま、 P は正定値とは限らない関数であるとする（複素数値を取っても良い）。 $P \propto e^{-S}$ と表されている場合、 S に虚部があると言ってもよい。そのような P は一般に

$$P(x_1, x_2, \dots, x_n) \propto e^{-\text{Re } S(x_1, x_2, \dots, x_n)} e^{-i \text{Im } S(x_1, x_2, \dots, x_n)} \quad (14)$$

と表すことができる。これを用いて積分 I を変形していくと、

$$I = \frac{\int dx_1 \dots dx_n O(x_1, \dots, x_n) e^{-S(x_1, \dots, x_n)}}{\int dx_1 \dots dx_n e^{-S(x_1, \dots, x_n)}} \quad (15)$$

$$= \frac{\int dx_1 \dots dx_n O(x_1, \dots, x_n) e^{-\text{Re } S(x_1, \dots, x_n)} e^{-i \text{Im } S(x_1, \dots, x_n)}}{\int dx_1 \dots dx_n e^{-\text{Re } S(x_1, \dots, x_n)} e^{-i \text{Im } S(x_1, \dots, x_n)}} \quad (16)$$

$$= \frac{\int dx_1 \dots dx_n O(x_1, \dots, x_n) e^{-\text{Re } S(x_1, \dots, x_n)} e^{-i \text{Im } S(x_1, \dots, x_n)} / \int dx_1 \dots dx_n e^{-\text{Re } S(x_1, \dots, x_n)}}{\int dx_1 \dots dx_n e^{-\text{Re } S(x_1, \dots, x_n)} e^{-i \text{Im } S(x_1, \dots, x_n)} / \int dx_1 \dots dx_n e^{-\text{Re } S(x_1, \dots, x_n)}} \quad (17)$$

となる。重み関数 P の下での期待値を $\langle \dots \rangle_P$ と書くことにすれば、

$$I = \langle O \rangle_{e^{-S}} = \frac{\langle O e^{-i \text{Im } S} \rangle_{e^{-\text{Re } S}}}{\langle e^{-i \text{Im } S} \rangle_{e^{-\text{Re } S}}} \quad (18)$$

が成り立つ。右辺は e^{-S} の代わりに $e^{-\text{Re } S}$ が重み関数として登場しているが、これは正定値なので確率密度の資格を持ち、モンテカルロ法が適用できる。このような書き換えを再重み付け (reweighting) と呼ぶ。

この結果は一見、モンテカルロ法は任意の P に適用できることを示しているように見える。しかしながら、 P の性質によっては

$$\langle e^{-i \text{Im } S} \rangle_{e^{-\text{Re } S}}, \langle O e^{-i \text{Im } S} \rangle_{e^{-\text{Re } S}} \sim e^{-V(n)} \quad (19)$$

のように振る舞う場合があることが知られている。ここで $V(n)$ 積分変数の数に依存して増大するなんらかの関数である。すると、これらの期待値を精度良く計算するためには

$$e^{-V} \gg \frac{1}{\sqrt{N}} \quad (20)$$

$$e^{2V} \ll N \quad (21)$$

のように指数的に大きなサンプル数 N が必要になってしまう。この現象は、 $e^{-i \text{Im } S}$ が激しい振動関数（正負に激しく揺れる）の場合に発生するので、符号問題 (sign problem) と呼ばれている。物理の文脈では、符号問題は相転移点の近傍で起きることが多い。

4 符号問題を回避する方法

符号問題の厳しさは $\langle e^{-i \text{Im } S} \rangle_{e^{-\text{Re } S}}$ がどれくらい小さいか、という指標 (phase factor) で特徴づけることができる。もしこれが 1 に十分近ければ、再重み付け法によって計算することが可能である。一方 phase factor が 0 に近い場合は、なんらかの対策が必要である。残念ながら、符号問題を完全に解決する方法は知られていない。ここでは、ある程度汎用性のある対策をいくつか挙げる*3。

4.1 テイラー展開法

確率分布 P がパラメータ α に依存していて、 $\alpha = \alpha_0$ では符号問題が発生しないとする。物理量 $O(\alpha)$ を α_0 の周りで展開すると、

$$O(\alpha) = \sum_n \frac{1}{n!} \frac{\partial^n O}{\partial \alpha^n} \Big|_{\alpha=\alpha_0} (\alpha - \alpha_0)^n \quad (22)$$

となる。仮定から、 $\frac{\partial^n O}{\partial \alpha^n} \Big|_{\alpha=\alpha_0}$ は符号問題なく計算が可能である。これによって $O(\alpha)$ を得る方法をテイラー展開法と呼ぶ。この手法は当然ながら、テイラー展開の収束半径内で有効であるが、問題を解かずして収束半径を知ることは困難なので、計算結果の信頼性をどう評価すべきかを、問題に応じて個別に検討しなくてはならない。

4.2 解析接続

確率分布 P がパラメータ α に依存していて、複素 α 平面上のある領域において符号問題が発生しないとする。そのような領域で $O(\alpha)$ を求めておき、領域外へ解析接続することができれば、符号問題を回避することができる。ただし、普通は領域内の離散的な点上において $O(\alpha)$ が求まるに過ぎないので、領域外への外挿には大きな不定性が生じることが多い。

*3 符号問題は有限密度量子色力学の文脈で詳しく研究されていて、[日本語のレビュー](#) もある。

4.3 複素ランジュバン法

q_i を n 自由度の力学変数とする。Langevin 方程式

$$\frac{dq_i}{dt} = K_i(q) + \eta_i(t) \quad (23)$$

を考える。 $K(q)$ は駆動項である。 $\eta_i(t)$ は

$$\langle \eta_i(t) \rangle = 0 \quad (24)$$

$$\langle \eta_i(t) \eta_j(s) \rangle = 2\alpha \delta_{ij} \delta(t-s) \quad (25)$$

を満たす Gaussian ノイズである。あるいは、 $\eta_i(t)$ は次の確率分布に従う確率変数である。

$$W_q(x)dx = \prod_i \frac{1}{\sqrt{4\pi\alpha}} \exp\left(-\frac{x_i^2}{4\alpha}\right) dx_i = \frac{1}{(4\pi\alpha)^{n/2}} \exp\left(-\frac{\sum_i x_i^2}{4\alpha}\right) dx \quad (26)$$

$\langle \dots \rangle$ はこの分布に関する統計平均である。伊藤型の離散化方程式は

$$q_i(t+dt) = q_i(t) + K_i(q)dt + dw_i(t) \quad (27)$$

で、ノイズは、

$$\langle dw_i(t) \rangle = 0 \quad (28)$$

$$\langle dw_i(t) dw_j(s) \rangle = \begin{cases} 0, & t \neq s \\ 2\alpha \delta_{ij} dt, & t = s \end{cases} \quad (29)$$

で与えられる。ラフに言って、 $\eta_i = dw_i/dt$ である。対応する確率分布関数は、

$$W_{dw}(x)dx = \frac{1}{(4\pi\alpha dt)^{n/2}} \exp\left(-\frac{\sum_i x_i^2}{4\alpha dt}\right) dx \quad (30)$$

である。離散化された Langevin 方程式のノイズ項のオーダーを明確にするため、ノイズ相関が時間刻みに依存しないようにリスケールすることが多い。すなわち、 $dw_i \rightarrow dw\sqrt{dt}$ として、

$$q_i(t+dt) = q_i(t) + K_i(q)dt + dw_i(t)\sqrt{dt} \quad (31)$$

で、ノイズは、

$$\langle dw_i(t) \rangle = 0 \quad (32)$$

$$\langle dw_i(t) dw_j(s) \rangle = \begin{cases} 0, & t \neq s \\ 2\alpha \delta_{ij}, & t = s \end{cases} \quad (33)$$

で与えられる。適当なテスト関数 $G(q)$ に対して、

$$\langle G(q(t)) \rangle = \int G(q) \Phi(q, t) dq \quad (34)$$

によって、確率分布関数 $\Phi(q, t)$ を定義する。 $\Phi(q, t)$ は次の Fokker-Planck 方程式に従うことが示せる。

$$\frac{\partial}{\partial t} \Phi(q, t) = \left[-\frac{\partial}{\partial q_i} K_i + \alpha \frac{\partial^2}{\partial q_i \partial q_i} \right] \Phi(q, t) \quad (35)$$

Fokker-Planck 方程式の解で

$$\frac{\partial}{\partial t} \Phi_{\text{eq}}(q) = 0 \quad (36)$$

を満たすものを定常解という。これは

$$\frac{\partial}{\partial q_i} \left[-K_i(q) + \alpha \frac{\partial}{\partial q_i} \right] \Phi_{\text{eq}}(q) = 0 \quad (37)$$

の解である。今、駆動項が何か（ポテンシャル）の微分で与えられているとする。

$$K_i(q) = -\frac{1}{f} \frac{\partial V(q)}{\partial q_i} \quad (38)$$

f は散逸を特徴づける定数であり、揺動散逸定理から $\alpha f = kT$ を満たす。すると、

$$\frac{\partial}{\partial q_i} \left[\frac{1}{f} \frac{\partial V(q)}{\partial q_i} + \alpha \frac{\partial}{\partial q_i} \right] \Phi_{\text{eq}}(q) = 0 \quad (39)$$

$$\frac{\partial}{\partial q_i} \left[\frac{\partial V(q)}{\partial q_i} + kT \frac{\partial}{\partial q_i} \right] \Phi_{\text{eq}}(q) = 0 \quad (40)$$

となる。もし、

$$\frac{\partial}{\partial q_i} \Phi_{\text{eq}}(q) = -\frac{1}{kT} \frac{\partial V(q)}{\partial q_i} \Phi_{\text{eq}}(q) \quad (41)$$

ならば、これは定常解である。よって、 C を規格化条件から定まる定数として、

$$\Phi_{\text{eq}}(q) = C \exp \left(-\frac{V(q)}{kT} \right) \quad (42)$$

は Fokker-Planck 方程式の定常解である。

これまでの議論で、

$$K_i(q) = -\frac{1}{f} \frac{\partial V}{\partial q_i} \rightarrow -\frac{\delta S[q]}{\delta q_i} \bigg|_{q=q(x,t)} \quad (43)$$

$$\alpha \rightarrow \hbar \quad (44)$$

とすると、Fokker-Planck 方程式の定常分布は、

$$\Phi(q, t) \propto \exp \left(-\frac{V(q)}{kT} \right) = \exp \left(-\frac{V(q)}{f\alpha} \right) \rightarrow \exp \left(-\frac{S[q]}{\hbar} \right) \quad (45)$$

となる。対応する Langevin 方程式は

$$\frac{dq_i}{dt} = -\frac{\delta S[q]}{\delta q_i} + \eta_i(x, t) \quad (46)$$

$$\langle \eta_i(x, t) \rangle = 0 \quad (47)$$

$$\langle \eta_i(x, t) \eta_j(y, s) \rangle = 2\hbar \delta_{ij} \delta^{(4)}(x - y) \delta(t - s) \quad (48)$$

である。よって、期待値を

$$\frac{\int dq O(q) e^{-S[q]}}{\int dq e^{-S[q]}} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{t_0}^T \langle O(q(t)) \rangle dt \quad (49)$$

によって計算することができる。この手続きは、一定の条件のもとで S が複素数の場合にも適用できる。これを複素ランジュバン法と呼ぶ。

4.4 Lefschetz thimble

簡単のため

$$Z = \int dx e^{-S(x)} \quad (50)$$

という一次元積分（分配関数と呼ぶ。）を考える。複素積分の範囲に拡張して考えると、この積分の積分経路は変更することができる。良い性質をもつ積分経路のひとつに Lefschetz thimble と呼ばれるものが知られている。作用 $S(x)$ を複素化したものを $S(z)$ としたとき、 $\partial S / \partial z = 0$ を満たすような点を臨界点、あるいは固定点と呼ぶ。次の flow 方程式

$$\frac{dz}{dt} = + \frac{\overline{\partial S}}{\partial z} \quad (51)$$

で定まる曲線のうち、固定点を通るようなものを Lefschetz thimble と呼ぶ。固定点を z_σ とし、対応する Lefschetz thimble を \mathcal{J}_σ と書く。このとき、分配関数は以下のように分解することができる。

$$Z = \int dx e^{-S(x)} = \sum_\sigma n_\sigma \int_{\mathcal{J}_\sigma} dz e^{-S(z)} \quad (52)$$

ここで n_σ は、もとの積分経路と次の flow 方程式

$$\frac{dz}{dt} = - \frac{\overline{\partial S}}{\partial z} \quad (53)$$

で定まる曲線の交差数を表す。この曲線を anti-thimble と呼び、 \mathcal{K}_σ と書く。

各 thimble 上で、複素作用がどのように変化するかを調べる。作用の虚部の変化は

$$\frac{d}{dt} \text{Im} S(z) = \frac{d}{dt} \frac{1}{2i} (S(z) - \overline{S(z)}) \quad (54)$$

$$= \frac{d}{dt} \frac{1}{2i} (S(z) - S(\bar{z})) \quad (55)$$

$$= \frac{1}{2i} \left(\frac{d}{dt} S(z) - \frac{d}{dt} S(\bar{z}) \right) \quad (56)$$

$$= \frac{1}{2i} \left(\frac{dz}{dt} \frac{\partial S(z)}{\partial z} - \frac{d\bar{z}}{dt} \frac{\partial S(\bar{z})}{\partial \bar{z}} \right) \quad (57)$$

$$= \frac{1}{2i} \left(\frac{dz}{dt} \frac{\partial S(z)}{\partial z} - \frac{d\bar{z}}{dt} \frac{\partial \overline{S(z)}}{\partial \bar{z}} \right) \quad (58)$$

$$= \frac{1}{2i} \left(- \frac{\overline{\partial S(z)}}{\partial z} \frac{\partial S(z)}{\partial z} + \frac{\partial S(z)}{\partial z} \frac{\overline{\partial S(z)}}{\partial z} \right) \quad (59)$$

$$= 0 \quad (60)$$

となる。作用の虚部が thimble 上で変化しないことから、その部分を積分の外に出すことができ、分配関数を

$$Z = \int dx e^{-S(x)} = \sum_{n_\sigma} e^{-\text{Im} S(z_\sigma)} \int_{\mathcal{J}_\sigma} dz e^{-\text{Re} S(z)} \quad (61)$$

と書くことができる。 J_σ 上で作用の実部は、

$$\frac{d}{dt} \operatorname{Re} S(z) = \frac{d}{dt} \frac{1}{2} (S(z) + \overline{S(z)}) \quad (62)$$

$$= \frac{d}{dt} \frac{1}{2} (S(z) + S(\bar{z})) \quad (63)$$

$$= \frac{1}{2} \left(\frac{d}{dt} S(z) + \frac{d}{dt} S(\bar{z}) \right) \quad (64)$$

$$= \frac{1}{2} \left(\frac{dz}{dt} \frac{\partial S(z)}{\partial z} + \frac{d\bar{z}}{dt} \frac{\partial S(\bar{z})}{\partial \bar{z}} \right) \quad (65)$$

$$= \frac{1}{2} \left(\frac{dz}{dt} \frac{\partial S(z)}{\partial z} + \frac{d\bar{z}}{dt} \overline{\frac{\partial S(z)}{\partial z}} \right) \quad (66)$$

$$= \frac{1}{2} \left(\frac{\partial \overline{S(z)}}{\partial \bar{z}} \frac{\partial S(z)}{\partial z} + \frac{\partial S(z)}{\partial z} \overline{\frac{\partial S(z)}{\partial z}} \right) \quad (67)$$

$$= \left| \frac{\partial S(z)}{\partial z} \right|^2 \geq 0 \quad (68)$$

のように変化する。まったく同様にして、 \mathcal{K}_σ 上では、

$$\frac{d}{dt} \operatorname{Re} S(z) = - \left| \frac{\partial S(z)}{\partial z} \right|^2 \leq 0 \quad (69)$$

のように変化する。したがって、固定点から離れるにつれて、 \mathcal{J}_σ 上では $e^{-\operatorname{Re} S(z)}$ が小さくなり、 \mathcal{K}_σ 上では大きくなる^{*4}。以上の議論から、thimble 上の積分は収束性が良く、かつ振動がないので、符号問題なく計算することが可能である。この手法は多重積分にもそのまま拡張可能である。ただし、thimble を具体的に求めるのは困難なので、応用上はもとの積分経路と thimble の間にある中間多様体上での積分を行うべきである。また、変数変換に伴うヤコビアン⁵の計算が重いことに留意すべきである。

5 量子計算のモンテカルロ法への応用

^{*4} flow 方程式の符号は人によって定義が異なるので注意。例えば、分配関数を

$$Z = \int dx e^{+S(x)} \quad (70)$$

のように作用の前の符号をプラスで定義した場合は、thimble を定める flow 方程式は

$$\frac{dz}{dt} = - \frac{\partial \overline{S}}{\partial z} \quad (71)$$

である。