

本レジュメの巻末には、Appendix のページを付けている。必要に応じて参照する。

## 1 ハイブリッドモンテカルロアルゴリズム

今回は Metropolis アルゴリズムを取り扱った。このアルゴリズムの制約として、状態空間を横切る移動距離がステップ数の平方根に比例する、逆に言うとステップ数が移動距離の二乗に比例するという、ランダムウォークの振る舞いを示してしまうことであった。ランダムウォークは状態空間の探索に時間がかかってしまい非効率である。

本節では、物理システムとの類似に基づいて、上記の問題を解決する手法を導く。そのために、まず今回は物理学の視点で力学系の枠組みを論じることにする\*1。そして、次回はその枠組みをもとに、Metropolis アルゴリズムとの組み合わせを行ない、ハイブリッドモンテカルロアルゴリズムを導出するという流れを踏む。

### 1.1 力学系

MCMC の目的は与えられた確率分布  $p(\mathbf{z})$  のサンプリングであった。この確率的なサンプリングをハミルトン系（のちに示すハミルトン方程式で記述される系）の形に当てはめるときに、ハミルトン力学の枠組みが必要となる。以下ではその枠組みの中身について詳しく説明していく。

まず力学系における位置変数  $z_i$  と運動量変数  $r_i$  の関係性を確かめよう。そのためにまず、ニュートンの第二法則、つまり運動方程式を思い出そう。

$$m \frac{d^2 z_i}{d\tau^2} = F \quad (1)$$

$m$  は質点の質量、 $\tau$  は時間、 $F$  は質点にかかる力である。ここで、さらに運動量が質量と速度の積\*2で書けたことも思い出してほしい。いま  $m = 1$  として無次元化してしまうと、速度を運動量として捉え直すことができる。 $r_i$  を  $r_i = \frac{dz_i}{d\tau}$  として約束すると、これは位置の一階微分、つまり速度に他ならないが、前文で「速度＝運動量」と述べたため、 $r_i$  は運動量として解釈することができるだろう。明らかに各位置変数と運動量変数とは一対一で対応していることがわかる。位置変数と運動量変数の結合空間を位相空間と呼ぶ\*3。

ここで確率分布  $p(\mathbf{z})$  を

$$p(\mathbf{z}) = \frac{1}{Z_p} \exp(-E(\mathbf{z})) \quad (2)$$

と表現しよう。 $E(\mathbf{z})$  はポテンシャルエネルギー\*4、 $Z_p$  は規格化定数である（はず）。ここまでの議論から、運動方程式は次のように書き直せる。

$$\begin{aligned} \frac{d}{d\tau} \frac{dz_i}{d\tau} &= -\frac{\partial E(\mathbf{z})}{\partial z_i} \\ \frac{dr_i}{d\tau} &= -\frac{\partial E(\mathbf{z})}{\partial r_i} \end{aligned} \quad (3)$$

\*1 教科書にも記載の通り、本節では物理学の背景知識は必要ない（勿論あった方が理解は深まる）。しかしながら、物理学で頻繁に用いるベクトル解析の知識は必須なので、こちらについては補足を適宜加えることにする。また、話は変わるが、教科書にある「第一原理」の意味は「実験事実や推論に依らない」ということである。ここでは、もう少し具体的に言うと、力学の基本法則である運動方程式のことを指していると解釈して良いだろう。

\*2 古典力学の範囲であるという前提はある。

\*3 物理屋は「相空間」と呼ぶ風習がある。

\*4 例えばバネポテンシャル  $\frac{1}{2}kx^2$  などがいたことを思い出そう。 $k$  はバネ定数、 $x$  は自然長からの伸び。

質点にかかる力（右辺）はポテンシャルの勾配に負号をつけたものとして約束している。なぜ負号をつけるのかは App を参照されたい。ここまではニュートン力学<sup>\*5</sup>の枠組みである。

ここからはハミルトン力学<sup>\*6</sup>の枠組みを用いて定式化をおこなう。まず運動エネルギーを

$$K(\mathbf{r}) = \frac{1}{2} \|\mathbf{r}\|^2 = \frac{1}{2} \sum_i r_i^2 \quad (4)$$

で定義する。このときハミルトニアン（ハミルトン関数）は系のエネルギーの総和<sup>\*7</sup>で表されて

$$H(\mathbf{z}, \mathbf{r}) = E(\mathbf{z}) + K(\mathbf{r}) \quad (5)$$

となる。この力学系は次のようなハミルトン方程式で表現することができる。

$$\frac{dz_i}{d\tau} = \frac{\partial H}{\partial r_i} \quad (6)$$

$$\frac{dr_i}{d\tau} = -\frac{\partial H}{\partial z_i} \quad (7)$$

以下では、それぞれの式を導くことにする。まず位置の時間微分に関する式は

$$\begin{aligned} \frac{\partial H}{\partial r_i} &= \frac{\partial}{\partial r_i} (E(\mathbf{z}) + K(\mathbf{r})) \\ &= \frac{\partial K(\mathbf{r})}{\partial r_i} \\ &= \frac{\partial}{\partial r_i} \frac{1}{2} \sum_i r_i^2 \\ &= \frac{1}{2} \cdot 2r_i \\ &= r_i \\ &= \frac{dz_i}{d\tau} \end{aligned} \quad (8)$$

となる。次に運動量の時間微分に関する式は

$$\begin{aligned} -\frac{\partial H}{\partial z_i} &= -\frac{\partial}{\partial z_i} (E(\mathbf{z}) + K(\mathbf{r})) \\ &= -\frac{\partial E(\mathbf{z})}{\partial z_i} \\ &= \frac{dr_i}{d\tau} \end{aligned} \quad (9)$$

となる。以上がハミルトン形式の定式化である。

ここからはハミルトン力学系において位相空間の同時分布が不変であることを導く。そのために次の 2 つの事実を示す必要がある。

<sup>\*5</sup> 初等力学のこと。高校で勉強した古典力学のことと思えばよろしい。

<sup>\*6</sup> なぜハミルトン形式を導入するのか。それはシンプルに言うと解析的な取り扱いが楽になるからである。例えば、ニュートン力学では、質点にかかる力について、何がどの向きでかかるのか考察をおこなった上で、運動方程式を立てる手続きを踏む。ニュートン力学を抜け出し、ハミルトン形式を導入すると、先のような考察のステップが排除される。ここで、次のような疑問「ハミルトニアンじゃなくてラグランジアンでも良いんじゃないの？」が湧いてくるかもしれない。この疑問に対する回答は <https://eman-physics.net/analytic/hamilton.html> を参照してほしい。ちなみにラグランジアンを適切にルジャンドル変換してやることで、ハミルトニアンを得ることができる。

<sup>\*7</sup> 上の注釈に記したルジャンドル変換の操作を行えば確かめられる。ここではやらないが。

系の時間発展において

1. ハミルトニアンは一定
2. 位相空間の体積は一定<sup>\*8</sup>

以下ではそれぞれ 1 と 2 が成立することを確かめよう。

#### 1. ハミルトニアンは一定

ハミルトニアン  $H$  を時間  $\tau$  で微分するだけで示せる。

$$\begin{aligned}\frac{dH}{d\tau} &= \sum \left( \frac{\partial H}{\partial z_i} \frac{dz_i}{d\tau} + \frac{\partial H}{\partial r_i} \frac{dr_i}{d\tau} \right) (\because \text{Chain rule}) \\ &= \sum \left( \frac{\partial H}{\partial z_i} \frac{\partial H}{\partial r_i} - \frac{\partial H}{\partial r_i} \frac{\partial H}{\partial z_i} \right) (\because \text{式 11.58-59}) \\ &= 0\end{aligned}\tag{10}$$

#### 2. 位相空間の体積は一定

流れ場（流体の速度を場の形で表したもの<sup>\*9</sup>）のダイバージェンス<sup>\*10</sup>が 0 になることを示す。流れ場は

$$\mathbf{V} = \left( \frac{dz}{d\tau}, \frac{d\mathbf{r}}{d\tau} \right)\tag{11}$$

であり、そのダイバージェンスは

$$\begin{aligned}\text{div} \mathbf{V} &= \left( \frac{\partial}{\partial z}, \frac{\partial}{\partial \mathbf{r}} \right) \cdot \left( \frac{dz}{d\tau}, \frac{d\mathbf{r}}{d\tau} \right) \\ &= \sum_i \left( \frac{\partial}{\partial z_i} \frac{dz_i}{d\tau} + \frac{\partial}{\partial r_i} \frac{dr_i}{d\tau} \right) \\ &= \sum_i \left( \frac{\partial}{\partial z_i} \frac{\partial H}{\partial r_i} - \frac{\partial}{\partial r_i} \frac{\partial H}{\partial z_i} \right) (\because \text{式 11.58-59}) \\ &= 0\end{aligned}\tag{12}$$

1. と 2. の結果を用いると、同時分布  $p(\mathbf{z}, \mathbf{r}) = \frac{1}{Z_H} \exp(-H(\mathbf{z}, \mathbf{r}))$  が不変となることがわかる。その説明は読み合わせ。

ここからはハミルトン力学系において時間発展をさせるときのサンプリングスキームについて考える。「いま  $H$  は一定であるので、 $p(\mathbf{z}, \mathbf{r})$  からのサンプリングがエルゴード的でない」この論理がよくわからなかった。エルゴード的なサンプリングスキームを得るために  $p(\mathbf{z}, \mathbf{r})$  を不変に保ち、 $H$  の値を変える方法を考える。これはギブスサンプリング<sup>\*11</sup>を用いることで実現できる。いま  $\mathbf{r}$  の値を  $\mathbf{z}$  で条件付けた分布から抽出したもので置き換えよう。 $p(\mathbf{r} | \mathbf{z})$  は、以下で確かめるようにガウス分布であり、容易にサンプリングを行なう

<sup>\*8</sup> リウヴィルの定理として知られている。

<sup>\*9</sup> 詳細は <https://eman-physics.net/fluid/flow.html> を参照。

<sup>\*10</sup> 発散、湧き出しとも呼ばれる。その概念は <http://hooktail.sub.jp/vectoranalysis/vecFuntou2/> を参照。

<sup>\*11</sup> ギブスサンプリングを使うと分布が不変となることは p.258 中段を見よ。

ことができる。

$$\begin{aligned}
p(\mathbf{r} \mid \mathbf{z}) &= \frac{p(\mathbf{z}, \mathbf{r})}{p(\mathbf{z})} \\
&= \frac{\frac{1}{Z_H} \exp(-H(\mathbf{z}, \mathbf{r}))}{\frac{1}{Z_P} \exp(-E(\mathbf{z}))} \quad (\because \text{式 11.54, 11.63}) \\
&= \frac{Z_P \exp(-E(\mathbf{z})) \exp(-K(\mathbf{r}))}{Z_H \exp(-E(\mathbf{z}))} \quad (\because \text{式 11.57}) \\
&= \frac{Z_P}{Z_H} \exp(-K(\mathbf{r})) \\
&= \frac{Z_P}{Z_H} \exp\left(-\frac{1}{2}\|\mathbf{r}\|^2\right) \quad (\because \text{式 11.56})
\end{aligned} \tag{13}$$

実際の応用、つまり微小時間のステップを動かして、注目している量<sup>\*12</sup>の差分を計算していくアプローチにおいては、必ず数値誤差が生じる<sup>\*13</sup>。したがって誤差を最小化する仕組みを考える必要があり、リープフロッグ離散化と呼ばれる手法がそれを実現する。この手法では、位置変数と運動量変数の離散時間近似  $\hat{\mathbf{z}}$  と  $\hat{\mathbf{r}}$  を式 (11.64-66) を用いて交互に更新する。アルゴリズムの詳細については読み合わせ。ちなみに本アルゴリズムではステップ  $\epsilon$  を二次まで拾う<sup>\*14</sup>ため、オイラー法などに比べれば誤差を抑えることができる。

---

<sup>\*12</sup> 今回は位置変数、運動量変数のこと。

<sup>\*13</sup> 例えばオイラー法では、ステップ  $\epsilon$  を一次まで拾うため、粗い近似となる。しかもオイラー法の場合は、その生じた誤差がステップを経る中で累積されていくため、トータルの誤差が大きくなるという問題が発生する。

<sup>\*14</sup> 式 (11.65) を (11.66) 左辺第二項に代入すれば明らかである。



# App ポテンシャルの勾配に負号を付ける理由

質量に与える力  $F$  とポテンシャル  $E$  の勾配に負号を付けたものとして約束した。

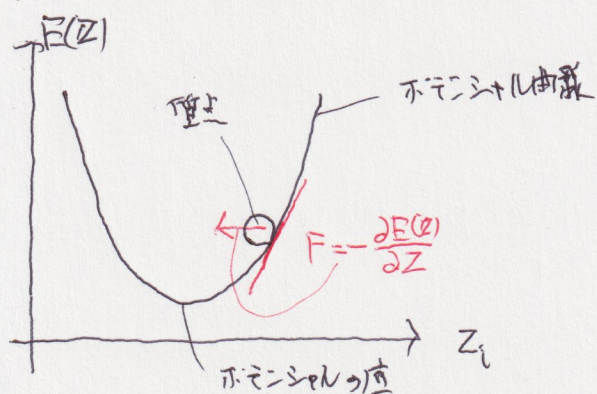
$$F_i = - \frac{\partial E(z)}{\partial z_i}$$

議論を簡単にするため 1 方向の成分のみ考えることにする。

負号を付ける理由は次の通りである：

「ポテンシャル場に存在する質点のふるまいと数式との関係を自然なものと解釈するため」

1D-シミュレーションを使って理解を深めよう。



左図のような状況で質点の動きも考える場合の自然な感覚として

「質点がポテンシャルの底に向かう」

ということが言えるだろう。

いま質点が存在する点におけるポテンシャル曲線の勾配は  $\frac{\partial E(z)}{\partial z}$  で

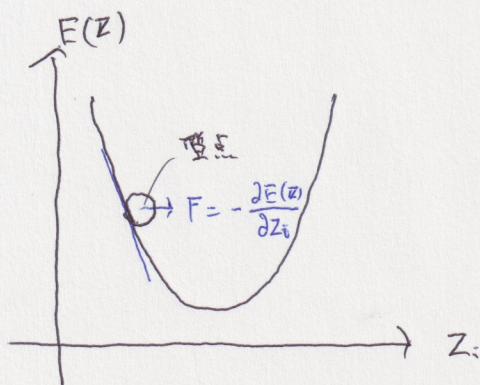
正の値になることがわかる。

→ この  $\frac{\partial E(z)}{\partial z}$  に負号を付けることの向きは図の通り左向きになる。

これはこの質点がポテンシャルの底に向かって運動するという

初めに述べた感覚に一致している。負号を付けることで

この感覚と数式の一致を体験することができる。



もちろん、左図のように質点がポテンシャルの底の左側にあるときも感覚と数式の一致が成り立つ。

注目している点の勾配は  $\frac{\partial E(z)}{\partial z}$  で負の値。

これに負号を付けること、力の向きは右向きでそれはこの質点がポテンシャルの底に向かうことを意味している。