

§ 4.3 Kinematic Interpretation

伊 冉 (Andre YI)

2026 年 1 月 22 日



PDF 版はここ ↑

1 概要

最適化問題における「関数の最小値を探索する」というプロセスは、物理学において「谷（ポテンシャル関数の形状）の中に置かれたボールが、重力に従って底まで転がり落ちる運動」として解釈することができます。本節では、質量 $m = 1$ のボールの運動を解析することで、最急降下法の物理的な意味と、単純な物理モデル（摩擦なし）では収束しない理由を数学的に解明する。

2 運動方程式の導出（ラグランジュ形式）

ボールの状態は、位置 x と速度 v のペア $s = (x, v)$ によって記述される。これを相空間 (phase space)¹ と呼ぶ。また、ボールには以下の 2 つのエネルギーが作用する。

- 運動エネルギー (Kinetic Energy): $E_k = \frac{1}{2}\|v\|^2$ （※質量 $m = 1$ のため）
- 位置エネルギー (Potential Energy): $E_p = f(x)$ （※目的関数の高さ）

物理学におけるラグランジュ量 (Lagrangian) L は、運動エネルギーと位置エネルギーの差として定義されます。

$$L(x, \dot{x}) = E_k - E_p = \frac{1}{2}\|\dot{x}\|^2 - f(x) \quad [1]$$

ここで、速度 v は位置の時間微分 \dot{x} と等しいことに注意してください。

Ex 1: Euler-Lagrangian 方程式による運動方程式の導出

ラグランジュ量 L を用いた Euler-Lagrangian 方程式：

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) = \frac{\partial L}{\partial x}$$

を用いて、ボールの運動方程式 $\ddot{x}(t) = -\nabla f(x(t))$ を導出しなさい。

Proof.

式 [1] を各項について偏微分する

最初は左辺の計算（運動量項） L を \dot{x} で偏微分する。 $f(x)$ は \dot{x} に依存しない定数項として扱う

¹ 相空間とは、全ての可能な状態の集合を指す物理学の用語。

れる。

$$\frac{\partial L}{\partial \dot{x}} = \frac{\partial}{\partial \dot{x}} \left(\frac{1}{2} \dot{x}^T \dot{x} \right) = \dot{x} = v$$

これをさらに時間 t で微分すると

$$\frac{d}{dt} v = \dot{v} = {}^2\ddot{x}$$

次に、右辺の計算（力項） L を x で偏微分する。 $\|\dot{x}\|^2$ は x に依存しないため消える。

$$\frac{\partial L}{\partial x} = -\frac{\partial}{\partial x} f(x) = -\nabla f(x)$$

これらを等号で結ぶと、以下のようになる。

$$\ddot{x}(t) = -\nabla f(x(t)) \quad [2]$$

これは、質量 $m = 1$ 、力 $F = -\nabla f(x)$ とした場合のニュートンの運動方程式 ($F = ma$) そのものである。 ■

3 エネルギー保存則とハミルトニアン

次に、運動中の「総エネルギー」の変化を考える。総エネルギー E_{tot} は、運動エネルギーと位置エネルギーの「和」である。

$$E_{tot}(t) = E_k(t) + E_p(t) = \frac{1}{2} \|\dot{x}(t)\|^2 + f(x(t)) \quad [3]$$

Ex 2: エネルギー保存則の証明

運動方程式 $\ddot{x} = -\nabla f(x)$ に従うとき、総エネルギー $E_{tot}(t)$ が時間変化しないこと、*i.e.* $\frac{d}{dt} E_{tot}(t) = 0$ を証明しなさい。

Proof.

総エネルギー式を時間 t で微分すると、以下のようになる。

$$\begin{aligned} \frac{d}{dt} E_{tot}(t) &= \frac{d}{dt} \left(\frac{1}{2} \dot{x}(t)^T \dot{x}(t) + f(x(t)) \right) \\ &= \frac{1}{2} \left(\left(\frac{d}{dt} \dot{x}(t) \right)^T \dot{x}(t) + \dot{x}(t)^T \frac{d}{dt} \dot{x}(t) \right) + \nabla f(x(t))^T \frac{d}{dt} (x(t)) \\ &= \dot{x}(t)^T \ddot{x}(t) + \dot{x}(t)^T \nabla f(x(t)) \\ &= \dot{x}(t)^T (\ddot{x}(t) + \nabla f(x(t))) \end{aligned}$$

² $\dot{v} = \frac{d}{dt} v = \frac{d}{dt} \left(\frac{d}{dt} x \right) = \frac{d^2}{dt^2} x = \ddot{x} = -\nabla f(x)$. ここで、 \ddot{x} は加速度であり、エックス・ツー・ドットと呼ぶ。

ここで、前節の [2] で導いた運動方程式より $\ddot{x}(t) + \nabla f(x(t)) = 0$ が成り立つ。したがって、

$$\frac{d}{dt} E_{tot}(t) = \dot{x}(t)^T \cdot 0 = 0$$

となり、エネルギーは保存されることが証明された。 ■

◆◆補足◆◆

この総エネルギーを相空間 (x, v) 上の関数として定義したものをハミルトニアン (**Hamiltonian**) H と呼ぶ。

$$H(x, v) = \frac{1}{2} \|v\|^2 + f(x)$$

エネルギーが保存されるということは、摩擦のないワンの中のボールは、永遠に転がり続け、決して底（最小値）に静止しないことを示唆している。

4 なぜ「摩擦なし」では収束しないのか

運動方程式は以下の 1 階連立常微分方程式 (ODE) 系として書き直せる。

$$\begin{cases} \dot{x}(t) = v(t) = \frac{\partial H}{\partial v} & [4] \\ \dot{v}(t) = -\nabla f(x(t)) = -\frac{\partial H}{\partial x} & [5] \end{cases}$$

この系におけるベクトル場の発散 (Divergence) を計算することで、収束性がわかる。

Ex 3: ベクトル場の発散の計算

相空間上のベクトル場 $X = (\dot{x}, \dot{v}) = \left(\frac{\partial H}{\partial v}, -\frac{\partial H}{\partial x} \right)$ の発散 $\text{div} X$ (i.e. $\nabla \cdot X$) を計算しなさい。

解答

発散の定義に従い、計算すると。

$$\begin{aligned} \text{div} X &= \frac{\partial}{\partial x}(\dot{x}) + \frac{\partial}{\partial v}(\dot{v}) \\ &= \frac{\partial}{\partial x} \left(\frac{\partial H}{\partial v} \right) + \frac{\partial}{\partial v} \left(-\frac{\partial H}{\partial x} \right) \\ &= \frac{\partial^2 H}{\partial x \partial v} - \frac{\partial^2 H}{\partial v \partial x} \end{aligned}$$

滑らかな関数 H において偏微分の順序交換が可能 ($\frac{\partial^2 H}{\partial x \partial v} = \frac{\partial^2 H}{\partial v \partial x}$) であるため、

$$\text{div} X = 0$$

となる。

◆◆◆補足◆◆◆

$\text{div} X = 0$ は、この物理系の流れが非圧縮 (**incompressible**) であることを意味する。これは、相空間内の状態の集合（体積）が、時間発展しても縮小しないことを意味する。最小点（平衡点）への収束は、状態空間の体積が一点に収縮することと等しいであるが、非圧縮性によりこれが不可能である。したがって、単純な物理モデル（摩擦なし）では、最適解へ収束せず、その周りを永遠に振動し続ける。収束させるには摩擦項の導入が必要不可欠である。

Appendix A ハミルトニアン (Hamiltonian) に対する最急降下法

Ex 4: 補足

もし、ハミルトニアン $H(x, v)$ 自体を目的関数と見なし、最急降下法を適用するとどうなるかを確認する。ここで、平衡状態 $S^* = (x^*, v^*)$, 初期状態 $S^0 = (x^0, v^0)$ とする。

H の勾配は以下ようになる

$$\nabla_s H(s) = \left(\frac{\partial H}{\partial x}, \frac{\partial H}{\partial v} \right) = (\nabla f(x), v)$$

これを用いた更新式 $s_{n+1} = s_n - \delta \nabla_s H(s_n)$ は、成分ごとに書くと以下の通りである。

$$\begin{aligned} x^{n+1} &= x^n - \delta \nabla f(x^n) \\ v^{n+1} &= v^n - \delta v^n = (1 - \delta)v^n \end{aligned}$$

ここで、 v^n とは n 回目の反復のインデックスである。上の結果を解析すると

1. v^n の更新式は $v^{n+1} = (1 - \delta)v^n \implies v^n = (1 - \delta)^n v^0$ となり、 δ が小さければ $v^n \rightarrow v^* = 0$ ($n \rightarrow \infty$) *i.e.* 速度は消える。
2. x の更新式は $x^{n+1} = x^n - \delta \nabla f(x^n)$ となり、これは通常的最急降下法と全く同じである。

つまり、エネルギー関数に対して単純に最急降下法を行っても、「速度成分がなくなり、単なる（運動量のない）最急降下法に戻ってしまう」という結果になる。運動量（慣性）の効果を得るには、運動方程式自体に摩擦項を組み込むアプローチが必要である。*i.e.* $\text{div} < 0$ が必要である。