

ニューラルネットワークを用いた ハミルトン系の Average Vector Field 法の改良

Improvement of the average vector field method
for Hamiltonian systems using neural networks

Elena Celledoni¹⁾ Brynjulf Owren²⁾ Chong Shen³⁾ Baige Xu⁴⁾ 谷口隆晴⁵⁾

¹⁾Department of Mathematical Sciences (Norwegian University of Science and Technology, E-mail: elena.celledoni@ntnu.no)

²⁾Department of Mathematical Sciences (Norwegian University of Science and Technology, E-mail: brynjulf.owren@ntnu.no)

³⁾神戸大学大学院理学研究科 (〒 657-8501 神戸市灘区六甲台町 1-1, E-mail: 242s026s@stu.kobe-u.ac.jp)

⁴⁾神戸大学大学院理学研究科 (〒 657-8501 神戸市灘区六甲台町 1-1, E-mail: baigexu@stu.kobe-u.ac.jp)

⁵⁾神戸大学大学院理学研究科 (〒 657-8501 神戸市灘区六甲台町 1-1, E-mail: yaguchi@pearl.kobe-u.ac.jp)

This paper focuses on energy-preserving numerical integrators for Hamiltonian systems. The average vector field method is a typical discrete gradient method that is used for designing energy-preserving numerical integrators. We propose a method to modify the average vector field method by employing neural networks to make it more accurate while preserving the energy-conservation property. We also show a numerical example, in which the proposed method in fact derives a discrete gradient that outperforms the average vector field method.

Key Words : Discrete gradient, Hamiltonian systems, Energy conservation law, Neural networks

1. はじめに

微分方程式の数値解法は計算工学の基礎をなすものであり、定量的・定性的に優れた手法の開発は重要である。特に、幾何的数値積分法 [1] は、方程式の持つ構造を保つ数値積分法であり、特定の幾何構造を持つ方程式に対して、より安定に、高精度な数値解を導くことができる。ハミルトン系に対しては、シンプレクティック数値積分法およびエネルギー保存型数値積分法が、代表的な構造保存型数値解法である。エネルギー保存型数値解法は、通常、離散勾配法によって導出される [1]。離散勾配法のうち、平均ベクトル場 (AVF) 法 [2] および Itoh-Abe 法 [3] は広く用いられている手法である。シンプレクティック積分法 [4] は、ハミルトンフローのものシンプレクティック構造を保つ数値解法である。

ニューラルネットワークは、任意の関数を近似できる普遍近似能力を持つため、常微分方程式の研究に頻繁に利用される。例えば、ニューラル常微分方程式 (Neural ODE) [6] は、支配方程式を定める常微分方程式の右辺項をモデル化することで、観測データから現象をモデル化できる。物理現象の数値計算への応用としては、シンプレクティックネットワーク (SympNets) [7] は、シンプレクティック構造を保ちつつ、ハミルトン系の解作用素を学習する。また、数値積分法においては、Runge-Kutta 法の最適係数をニューラルネットワークを用いて求める方法も提案されている [5]。

本研究では、ニューラルネットワークを用いて平均ベクトル場 (AVF) 离散勾配法を修正する手法を提案する。これによって、エネルギー保存則を保ちつつ、真の解により近い数値解を求めることが可能となる。

Remark 1 本研究では、AVF 法に焦点を当てるため、

AVF 离散勾配を解析的に求めることが可能なハミルトン系を仮定する。他の離散勾配への応用については今後の課題とする。

2. 离散勾配法

Definition 1 ([8]) 滑らかな関数 $H : \mathbb{R}^d \rightarrow \mathbb{R}$ に対する離散勾配とは、任意の $x, y \in \mathbb{R}^d$ に対して、以下の性質を満たす写像 $\bar{\nabla}H : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ である：

$$\begin{cases} \bar{\nabla}H(x, y)^\top (y - x) = H(y) - H(x), \\ \bar{\nabla}H(x, x) = \nabla H(x). \end{cases}$$

第一の条件は本質的に離散勾配に対する離散版の連鎖律を表し、第二の条件は連続版の勾配の近似になっていることを保証する。

本論文では、簡単のため、以下の形のハミルトン系を考える：

$$\dot{x} = S \nabla H(x),$$

S は x に依存しない、可逆な歪対称行列であるとする。数値解を $x^{(n)} \approx x(nh)$ ($n = 0, 1, 2, \dots$) などと表すものとする。また、時間ステップ幅を h とする。離散勾配を用いた数値スキームは次のように定義される：

$$\frac{x^{(n+1)} - x^{(n)}}{h} = S \bar{\nabla}H(x^{(n)}, x^{(n+1)}).$$

離散勾配の定義より、上記の数値積分法は以下の意味

で離散エネルギー保存性を満たすことが示される：

$$\begin{aligned} H(x^{(n+1)}) - H(x^{(n)}) \\ = \bar{\nabla}H(x^{(n)}, x^{(n+1)})^\top (x^{(n+1)} - x^{(n)}) \\ = h\bar{\nabla}H(x^{(n)}, x^{(n+1)})^\top S\bar{\nabla}H(x^{(n)}, x^{(n+1)}) \\ = 0. \end{aligned}$$

本研究では、AVF法[2]に焦点を当てる。与えられたハミルトニアン H に対して、AVF離散勾配は次のように定義される：

$$\bar{\nabla}_{\text{AVF}}H(x, y) = \int_0^1 \nabla H((1-\xi)x + \xi y) d\xi.$$

また、AVF法による数値スキームは次のように表される：

$$\begin{aligned} \frac{x^{(n+1)} - x^{(n)}}{h} &= S\bar{\nabla}_{\text{AVF}}H(x^{(n)}, x^{(n+1)}) \\ &= S \int_0^1 \nabla H((1-\xi)x^{(n)} + \xi x^{(n+1)}) d\xi. \end{aligned}$$

AVF法はエネルギー保存則と時間対称性を保つ。しかし、AVF法の定式化には平均勾配の計算のための厳密な積分が必要であるため、多項式ハミルトニアンのような比較的簡単なハミルトニアンを考慮しない限り、実装が困難である。

3. AVF離散勾配の例

AVF離散勾配法を説明するために、単純な振り子を例として用いる。本論文で提案する手法は、AVF離散勾配を解析的に求めることができる限り、他のハミルトン系にも適用可能であることに留意する。

振り子は、質量なしの糸の一端を固定し、他端におもりを付けて吊るされ、重力の影響下で揺れを繰り返す系である。この系は、非線形ハミルトン系の例として用いられることがある。

質量 $m = 1$ 、糸の長さ $l = 1$ 、重力加速度 $g = 1$ の振り子は、以下のハミルトニアンをもつ：

$$H(q, p) = \frac{1}{2}p^2 - \cos(q).$$

ここで、位置座標 q は鉛直方向からの角度を表し、 $p = \dot{q}$ は運動量を表す。ベクトル $x = (q, p)^\top$ を導入すると、ハミルトン方程式は以下のように表される：

$$\dot{x} = S\nabla H(x), \quad S = \begin{pmatrix} O & I \\ -I & O \end{pmatrix}.$$

定義より、この系に対するAVF離散勾配は次のように表される：

$$\begin{aligned} \bar{\nabla}_{\text{AVF}}H(x^{(n)}, x^{(n+1)}) \\ = \int_0^1 \nabla H((1-\xi)x^{(n)} + \xi x^{(n+1)}) d\xi \\ = \begin{pmatrix} \int_0^1 \frac{\partial H}{\partial q}((1-\xi)x^{(n)} + \xi x^{(n+1)}) d\xi \\ \int_0^1 \frac{\partial H}{\partial p}((1-\xi)x^{(n)} + \xi x^{(n+1)}) d\xi \end{pmatrix}. \end{aligned}$$

上記の積分は解析的に求めることができる。具体的には

$$\begin{aligned} &\int_0^1 \frac{\partial H}{\partial q}((1-\xi)x^{(n)} + \xi x^{(n+1)}) d\xi \\ &= \int_0^1 \sin((1-\xi)q^{(n)} + \xi q^{(n+1)}) d\xi \\ &= \frac{-\cos(q^{(n+1)}) + \cos(q^{(n)})}{q^{(n+1)} - q^{(n)}} \end{aligned}$$

や

$$\begin{aligned} &\int_0^1 \frac{\partial H}{\partial p}((1-\xi)x^{(n)} + \xi x^{(n+1)}) d\xi \\ &= \int_0^1 (1-\xi)p^{(n)} + \xi p^{(n+1)} d\xi \\ &= \frac{1}{2}(p^{(n)} + p^{(n+1)}). \end{aligned}$$

となる。以上から、AVF離散勾配は次のように導出される：

$$\bar{\nabla}_{\text{AVF}}H(x^{(n)}, x^{(n+1)}) = \begin{pmatrix} -\cos(q^{(n+1)}) + \cos(q^{(n)}) \\ \frac{q^{(n+1)} - q^{(n)}}{\frac{1}{2}(p^{(n)} + p^{(n+1)})} \end{pmatrix}.$$

4. 提案手法

AVF法はエネルギー保存則と、時間対称性を保つものの、数値誤差が蓄積し時間とともに増大する可能性がある。本研究では、AVF法に基づく離散勾配をニューラルネットワークを用いて最適化し、エネルギー保存性を維持しつつ、数値解を真の解により近づけることを目的とする。さらに、本研究では固定時間ステップの場合のみを考慮するが、今後の研究では可変時間ステップの場合についても検討する予定である。

提案手法は次のように定式化される：

$$\bar{\nabla}_{\text{nn}}H(x^{(n)}, x^{(n+1)}) = \bar{\nabla}_{\text{AVF}}H(x^{(n)}, x^{(n+1)}) + M(x^{(n+1)} - x^{(n)}),$$

ここで、 $M(x^{(n)})$ は、 $x^{(n)}$ を入力とするニューラルネットワークによって学習される歪対称行列である。

この手法によって得られる新たな勾配 $\bar{\nabla}_{\text{nn}}H(x^{(n)}, x^{(n+1)})$ は、離散勾配の定義における2つの条件を満たすことが示される。したがって、提案する数値スキームはエネルギー保存性をもつ。

実際、まず、

$$\begin{aligned} &\bar{\nabla}_{\text{nn}}H(x^{(n)}, x^{(n+1)})^\top (x^{(n+1)} - x^{(n)}) \\ &= \bar{\nabla}_{\text{AVF}}H(x^{(n)}, x^{(n+1)})^\top (x^{(n+1)} - x^{(n)}) \\ &\quad + (x^{(n+1)} - x^{(n)})^\top M^\top (x^{(n+1)} - x^{(n)}) \\ &= \bar{\nabla}_{\text{AVF}}H(x^{(n)}, x^{(n+1)})^\top (x^{(n+1)} - x^{(n)}) \\ &= H(x^{(n+1)}) - H(x^{(n)}). \end{aligned}$$

が成り立ち、また、

$$\bar{\nabla}_{\text{nn}}H(x^{(n)}, x^{(n)}) = \bar{\nabla}_{\text{AVF}}H(x^{(n)}, x^{(n)}) = \nabla H(x^{(n)})$$

も成り立つ。

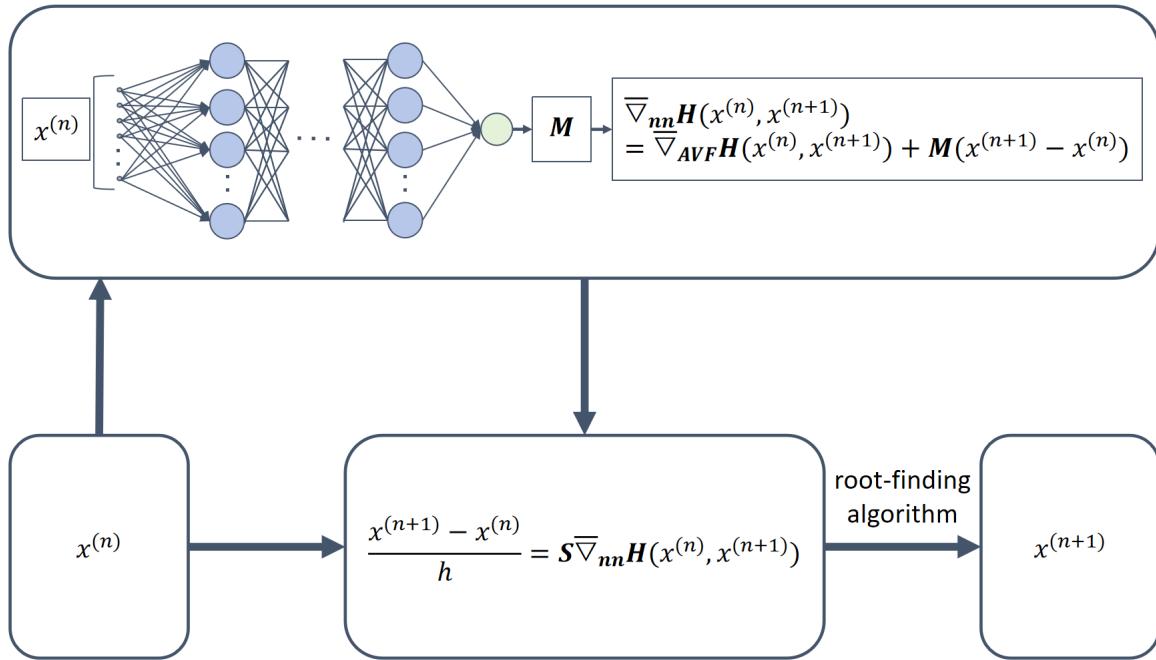


図-1 提案手法のアーキテクチャ

提案手法のアーキテクチャは図 1 に示されている。行列 M を生成するニューラルネットワークは、平均二乗誤差損失を最小化することによって学習される：

$$\mathcal{L} = \frac{1}{2N} \sum_{n=1}^N \left\| \frac{x^{(n+1)} - x^{(n)}}{h} - S \bar{\nabla}_{nn}H(x^{(n)}, x^{(n+1)}) \right\|^2.$$

ここで、 $\{(x^{(n)}, x^{(n+1)}) \mid n = 0, \dots, N\}$ は、正確に計算された数値解から構成されるトレーニングデータセットである。

5. 数値実験

提案手法の有効性を確認するために数値実験を行った。トレーニングデータは、固定時間ステップ幅 $h = 0.5$ で高次の数値積分法を用いて計算された軌道から得られた $N = 15$ 点から構成される。ただし、初期値は $x^{(0)} = (2, 0)$ とした。より正確には、データは、ある時間ステップ n での数値解と次の時間ステップでの数値解のペアで構成され、 $\{(x^{(n)}, x^{(n+1)}) \mid n = 0, \dots, N\}$ となる。図 2 はトレーニングデータを示している。ハイパーパラメータとして、学習率は 0.00001 に設定し、学習のエポック数は 30000 に設定した。

学習後、ニューラルネットワークで最適化された離散勾配を使用して、学習用データセットの終了点である $x^{(N+1)}$ をテストデータの初期値とし、そこから 200 個のテストデータを取った。図 3 に、学習した離散勾配を用いて計算された (q, p) の予測軌道を、図 4 に、時間に対する位置 q の発展を、図 5 に、時間に対する運動量 p の発展を示す。

提案した数値スキームはニューラルネットワークを使用するため、計算コストが AVF 法よりも高くなる。そのため、これを考慮して、計算時間も比較する。振

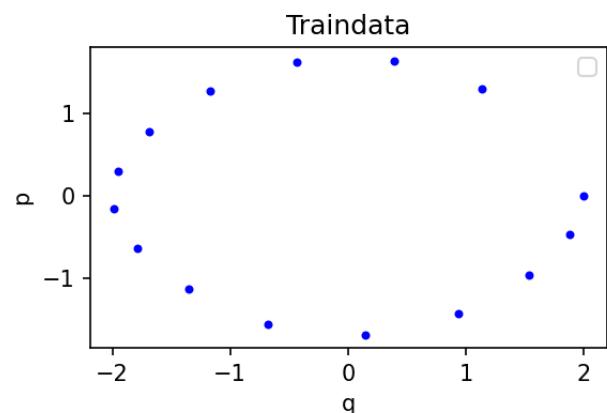


図-2 振り子の学習データ

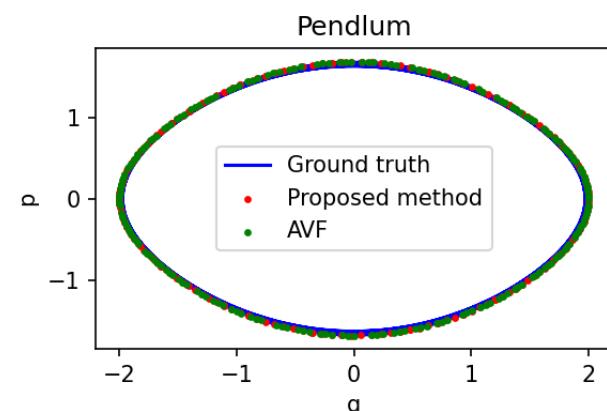


図-3 振り子のテストフロー

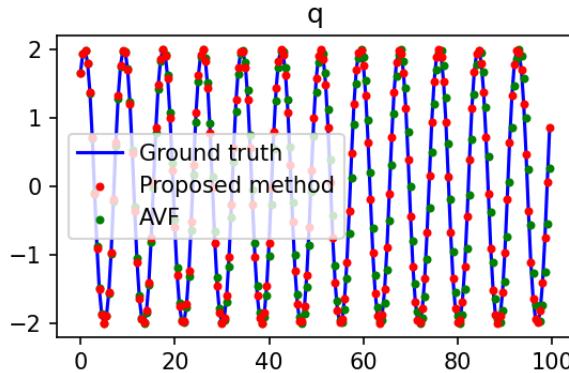
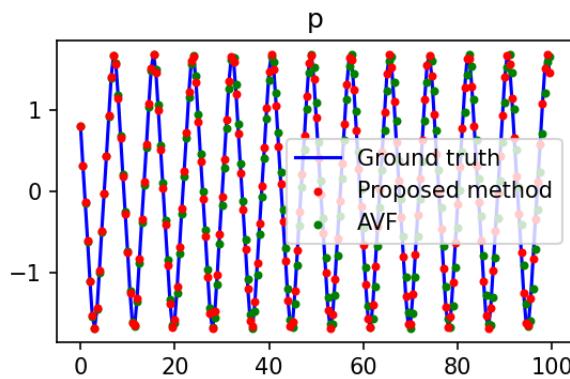
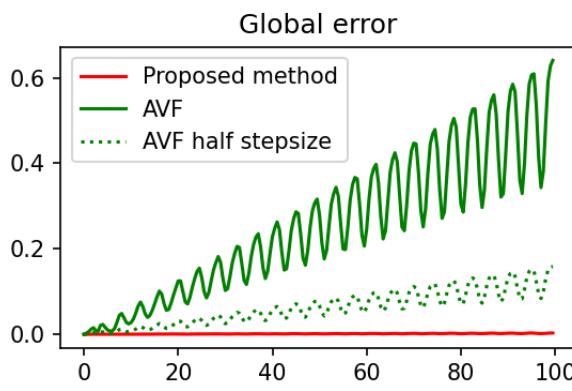
図-4 テストフローにおける位置 q 図-5 テストフローにおける運動量 p 

図-6 テストフローにおけるグローバル誤差

り子の実験で 5 層、幅 50 のネットワークを使用した場合、200 点の軌道計算にかかる計算時間は、AVF 法と、時間ステップサイズを半分にした AVF 法の間に位置していた。例えば、具体的な計算時間は、提案手法は 1.300 [秒]、AVF 法は 0.818 [秒]、半分のステップサイズの AVF 法は 1.392 [秒] などとなった。

図 6 に、予測された数値解と真の解の差のノルムである大域的誤差 $\|x^{(n)} - x(nh)\|_2$ を示す。また、図 7 はエネルギーの変化を示す。AVF 法および提案手法の両方

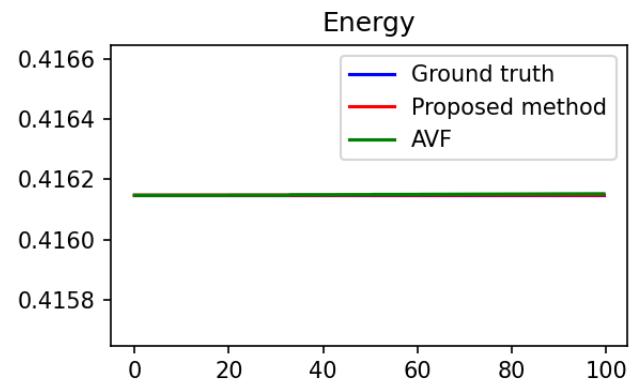


図-7 テストフローにおけるエネルギー

で、非線形方程式を解く必要があるため、誤差が発生する可能性があるが、結果からエネルギーの変動が、求解のためのアルゴリズムの許容誤差内に収まっていることが確認できる。

6. 結論

本研究では、AVF 離散勾配をニューラルネットワークを用いて修正し、数値スキームをより高精度にするための離散勾配の導出方法を提案した。また、数値実験によって、手法の有効性が確認できた。

しかしながら、まず本研究では提案手法を振り子以外のハミルトン系にも適用したものの、提案手法の実用性は、網羅的に確認できてはいない。したがって、提案手法が他のハミルトン系に対しても十分に実用的であるかを、さらに確認することが重要である。次に、本論文で使用した提案手法は AVF 法に基づいており、Itoh-Abe 法などの他の離散勾配法を使用した場合でも同様の改善が得られるかどうかを確認する必要がある。さらに、提案手法の可変時間ステップへの拡張についても検討することが重要である。これらの課題については、今後の研究で取り組んでいく予定である。

謝辞: 本研究は、JST CREST JPMJCR1914、JPMJCR24Q5 および JST ASPIRE JPMJAP2329 の支援を受けている。

参考文献

- [1] Hairer, E., Lubich, C., Wanner, G.: *Geometric numerical integration*, Springer, 2006.
- [2] Quispel, G. R. W., McLaren, D. I.: A new class of energy-preserving numerical integration methods, *J. Phys. A*, Vol. 41, 045206, 2008.
- [3] Itoh, T., Abe, K.: Hamiltonian-conserving discrete canonical equations based on variational difference quotients, *J. Comput. Phys.*, Vol. 76, pp. 85-102, 1988.
- [4] Feng, K., Qin, M.: *Symplectic geometric algorithms for Hamiltonian systems*, Springer, 2010.
- [5] Anastassi, A. A.: Constructing Runge–Kutta methods with the use of artificial neural networks., *Neural*

- Computing and Applications, *Neural Comput. Appl.*, Vol. 25, pp. 229-236, 2014.
- [6] Chen, R. T., Rubanova, Y., Bettencourt, J., Duvenaud, D. K.: Neural ordinary differential equations, *Adv. Neural Inf. Process. Syst.*, Vol. 31, pp. 6572-6583, 2018.
- [7] Jin, P., Zhang, Z., Zhu, A.: SympNets: Intrinsic structure-preserving symplectic networks for identifying Hamiltonian systems, *Neural Netw.*, Vol. 132, pp. 166-179, 2020.
- [8] Gonzalez, O.: Time integration and discrete Hamiltonian systems, *J. Nonlinear Sci.*, Vol. 6, pp. 449-467, 1996.