4A-02

ニューラルネットワークを用いたリーマン計量に対する変分問題の数値解法

佐藤 哲 †
パーソルキャリア株式会社 †

1. はじめに

ニューラルネットワークにより関数を近似し，ニューラルネットワークの微分と数値積分により微分方程式を解き，数値的に解を求める研究が盛んである[1][2][3][4]．ニューラルネットワークが近似する関数は変分問題で使われる積分の被積分関数であり，目的に応じて多くのものが提案されている．本研究では，計算コストを考慮しつつ色々な変分問題に対応可能なリーマン計量に基づく関数を提案する．

2. ニューラルネットワークを用いた変分問題の数値解法

変分問題は，関数 F を積分することによって定まる汎関数

$$S = \int_{t_0}^{t_1} F(t, x, x')dt. \tag{1}$$

の極値を与える関数の形を決定する問題である．ここで，t は独立変数，x は t の関数，x' は x の t に関する導関数 dx/dt である．変分問題の課題は，対象に応じた微分方程式を導くこと及び導いた微分方程式を解き軌道を求めることである．この課題に対する汎用的なアプローチが，ニューラルネットワークを用いた変分問題の数値解法である．そのアルゴリズムの概要は以下のようなものである．

(1) 関数 F の値を求めるネットワークを用意する
(2) 観測データを元にネットワークを訓練する．ただし，関数 F の値は観測できないため，F が満たすべき微分方程式ついての損失関数を元に学習する
(3) 学習後のネットワークを微分することにより微分方程式を導出し，微分方程式を数値的に解くことで座標や速度を求める

このアルゴリズムの特徴は，入力データから関数の特徴を学習し，任意の初期値による軌道が推定可能なことである．数学的には微分方程式を解き関数 F の形を決定し計算する必要があるが，関数 F をニューラルネットワークにより近似することができれば，関数の形を決定しなくても様々な計算が可能である．式（1）に基づき変分法により関数 F が満たすべき方程式を導き，方程式を解き座標データを求めるために，関数 F として座標成分及び運動量成分の関数であるハミルトニアンや，座標成分及び速度成分の関数であるラグランジアンなどが提案されている．本研究では，座標成分による関数及び速度成分の 2 次以下の項により構成されるリーマン計量に基づく関数を対象とする：

$$F = \sum_{i=0}^{n-1} g_i \left(\frac{dx_i}{dt}\right)^2 \tag{2}$$

ここで，g_i は座標成分 $(x_0, \cdots, x_{n-1})$ の関数であり計量テンソルと呼ばれる．式（1）において，S が極値を取る必要条件は次のオイラーラグランジュの方程式を満たすことであることが知られている：

$$\frac{\partial F}{\partial x} = \frac{d}{dt}\frac{\partial F}{\partial x'} \tag{3}$$

式（2），式（3）より，次式が得られる：

$$\sum_{i=0}^{n-1} \frac{\partial g_i}{\partial x_j}\left(\frac{dx_i}{dt}\right)^2 = 2\frac{dg_j}{dt}\frac{dx_j}{dt} + 2g_j\frac{d^2x_j}{dt^2} \tag{4}$$

ここで，$0 \leq j < n-1$ である．式（4）は関数 F を特徴付ける条件であり，この条件を満たす関数を学習するために，次の損失関数を最小化するようにニューラルネットワークを訓練する：

$$loss_j = \left\| \sum_{i=0}^{n-1} \frac{\partial g_i}{\partial x_j}\left(\frac{dx_i^*}{dt}\right)^2 - 2\frac{dg_j}{dt}\frac{dx_j^*}{dt} - 2g_j\frac{d^2x_j^*}{dt^2} \right\|_p \tag{5}$$

ここで，dx_i^*/dt 及び $d^2x_i^*/dt^2$ は学習データである．また，記号 $\|\|_p$ はノルムを表す．

式（4）は x の 2 階導関数について解くことができるため，として計量テンソル g_i 及び g_j を含む項が評価できれば，x の 1 階導関数の値を初期値として任意の数値解法により数値的に解き，解軌道 $\{x\}$ を求めることができる．本研究では，計量テンソルはニューラルネットワークの出力として求めることができ，ニューラルネットワークを微分可能なように構成することにより，導関数も計算可能だる．従って，データによりニューラルネットワークを訓練し，データが表す関数の任意の座標を求めるアルゴリズムは，以下のようになる．

(1) 計量テンソル $g_0, g_1, \cdots, g_{n-1}$ を求めるネットワークを用意する
(2) 観測データである速度，加速度により損失関数（5）を計算し，ネットワークを訓練する．
(3) 学習後のネットワークにより式（4）を構成し，数値的に解くことで座標や速度を求める

Numerical solutions of variational problems for Riemannian metrics by neural networks
†Tetsu R. Satoh, PERSOL CAREER CO., LTD.

Copyright ©2022 Information Processing Society of Japan.
All Rights Reserved.

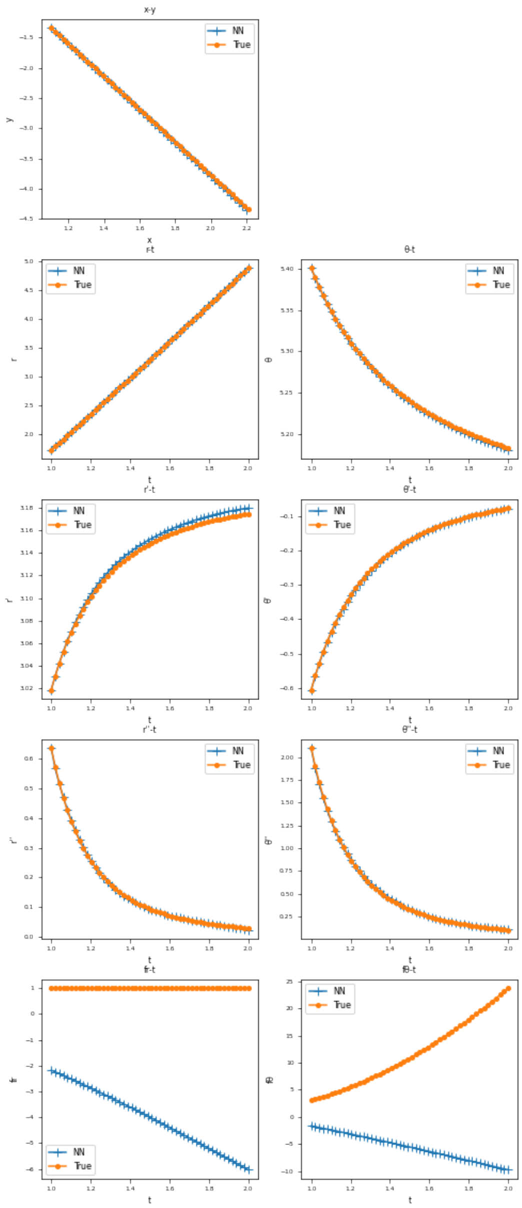

図 1: 極座標系の直線の計算結果

3. 実験例

簡単な例として，2 次元平面上の 2 点間の最短経路を求める変分問題について数値実験を行う．

平坦な 2 次元平面上では 2 点間の最短経路は直線であり，極座標系では

$$r\cos(\theta-\theta_0)=r_0 \tag{6}$$

と表される．直交座標系での直線とは違い，座標を独立変数の関数として陽に表すことができないため，数値的な処理が必要な問題である．

実験は，まず学習データとして式（6）より座標成分 (r,θ) の 1 階導関数の値をサンプリングして用意する．そして前節にて述べたアルゴリズムに従いニューラルネットワークを訓練し，訓練後のニューラルネットワークを利用して数値的に座標や速度を求めた．各パラメータをプロットした結果を図 1 に示す．図は，最上部より (1) 座標成分 r 及び θ から算出した x 座標及び y 座標，(2) 座標成分 r 及び θ，(3) 速度成分 r' 及び θ'，(4) 加速度成分 r'' 及び θ''，(5) 関数 g を表し，最上部の図を除き横軸は独立変数 t である．全ての図において丸のマーカーで示されるグラフは理論値であり，「+」のマーカーで示されるグラフはニューラルネットワークを利用して数値的に求めた結果である．

図より，数値計算の誤差が見られるものの，学習データより極座標系の座標成分の直線を特徴付けるパラメータを推定できていることが分かる．また，座標系の対角成分のみを対象としているために逆行列の計算が発生せず，一般的な条件と比べて計算量は削減されている．ただし，図 1 の最下部の図に示されるように，計量テンソルの成分については理論値を再現できていない．この問題は，従来研究のハミルトニアン・ニューラルネットワークやラグランジアン・ニューラルネットワークにおいてもハミルトニアンやラグランジアンの値が正確には再現できていない現象と同じであり，今後の検討が必要である．

本実験では検証のために精度を計りやすい直線を対象としたが，実験においては直線であることの情報は利用していなく，あらゆる入力データから特徴を学習して任意の初期値による軌道を推定可能であることが期待される．

以上の実験は Amazon Web Services(AWS) 上のインスタンス r5.xlarge にて，python3.7 及び torch1.9 などを利用して実施した．

4. おわりに

本研究では，変分原理の目的である解軌道を求めるために，ニューラルネットワークを利用してリーマン計量に基づく関数を対象として計算するアルゴリズムを提案し，実験により検証した．実験では 2 次元平面上で 2 点間の最短経路を求める問題に対し精度良く解軌道が得られることが示されたが，さらに多くの問題に対し実装を進め検証することが必要である．

参考文献

[1] S. Greydanus, M. Dzamba and J. Yosinski, Hamiltonian Neural Networks, *arXiv:1906.01563*, 2019.

[2] M. Cranmer, S. Greydanus, S. Hoyer, P. Battaglia, D. Spergel and S. Ho, Lagrangian Neural Networks, *arXiv:2003.04630*, 2020.

[3] T. Matsubara, A. Ishikawa and T. Yaguchi, Deep Energy-Based Modeling of Discrete-Time Physics, *arXiv:1905.08604*, 2020 .

[4] T. Aoshima, T. Matsubara and T. Yaguchi, Deep Discrete-Time Lagrangian Mechanics, ICLR2021 Workshop on Deep Learning for Simulation, 2021.

Copyright ©2022 Information Processing Society of Japan.
All Rights Reserved.