

有田拓朗*, 前田道治*
(*福岡工業大学大学院工学研究科)

1. はじめに

深層学習は画像認識や自然言語処理に加え、物理現象や時系列解析など連続データの予測にも応用が拡大している。予測には回帰と、未観測領域を推定する外挿があり、特に後者は未知の挙動を推定する必要があり困難である。本稿では、ノイズを含むスパイラルデータを対象にその外挿性能を検証する。

2. ニューラルネットワーク

従来の深層学習モデルとして広く用いられてきたのが多層パーセプトロン (MLP) や畳み込みニューラルネットワーク (CNN) である。これらは入力値を層ごとに非線形変換し特徴を抽出するが、深層化に伴い勾配消失や学習困難が問題となっていた。この課題を解決するために提案されたのが ResNet [1]であり、第 $t+1$ 層目の隠れ層の状態を h_{t+1} とすると、その基本構造は、式(1)のような残差接続によって表される。

$$h_{t+1} = h_t + f(h_t, \theta_t) \quad (1)$$

ここで、 f は学習される変換関数、 θ_t は学習可能なパラメータを表す。この仕組みにより勾配が層をまたいで伝播しやすくなり、非常に深いネットワークの学習が可能となった。しかし ResNet は依然として層の離散的な積み重ねに基づくため、連続的な力学系の記述や外挿予測には必ずしも適していないとされている。

3. ニューラル微分方程式

ニューラル微分方程式[2]は、ニューラルネットワークを微分方程式として定式化したモデルであり、従来の層構造を連続的に扱うことで柔軟で表現力の高いモデルを実現する。特に順方向、逆方向の伝播に特徴がある。順伝播では隠れ層のダイナミクスを式(2)のような微分方程式で記述し、その解を数値的に求めることで出力値と損失関数 L を計算する。

$$\frac{d}{dt}h(t) = f(h(t), t, \theta) \quad (2)$$

ここで、 $h(t)$ はある時間 t における隠れ状態を表す。逆伝播では、損失関数により求められた誤差を後ろから前に伝播し、最初の層のパラメータの勾配を求める必要がある。具体的には、adjoint methodという方法で式(3)のように目的の勾配を求めることができ、パラメータ更新がなされる。

$$\frac{\partial L}{\partial \theta} = - \int_{t_1}^{t_0} a(t)^T \frac{\partial f(z(t), t, \theta)}{\partial \theta} dt \quad (3)$$

ここで、 $z(t)$ はある時間 t における隠れ状態を表す。式(2)、式(3)で示したように、ニューラル微分方程式では、順伝播のみならず逆伝播も常微分方程式の数値解法により評価できることが分かる。この特徴が学習におけるメモリ効率の向上にも繋がっている。

4. 数値実験

本稿では、生成した観測点に対し、ニューラル微分方程式を用いて回帰及び外挿ができるか観測した。データセットは式(4)のような極座標の方程式から得られるスパイラルデータにノイズを加え、ばらつきのある観測点を生成した。ノイズの標準偏差は0.25、観測点は60点生成し、活性化関数には Tanh 関数を用いた。

$$x = a\theta \cos\theta, y = a\theta \sin\theta \quad (4)$$

ここで、 a は0より大きい定数であり、本実験では $a = 1$ に設定した。このとき等間隔な螺旋を描く。

図1に学習後の結果を示す。

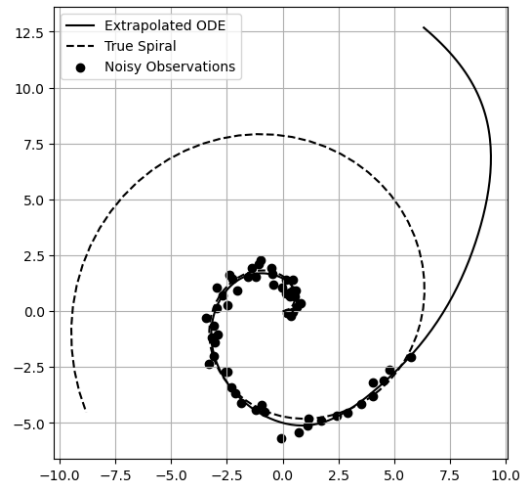


図1 実験結果

図1において、横軸と縦軸はそれぞれ式(4)の x 軸、 y 軸を表し、破線が式(4)によって生成した真のスパイラル曲線、黒点がスパイラル座標を元に生成したノイズ付き観測点、実線がニューラル微分方程式の外挿結果である。図1から、ばらつきのあるデータに対しては滑らかに回帰できている一方で、外挿部分については真のスパイラルから大きく外れていることが確認できる。

5. まとめ

本稿では、ニューラル微分方程式を用いてスパイラルデータの回帰及び外挿の実験を行った。今後の課題として、外挿精度の向上と実行時間の短縮を挙げられる。

参考文献

- [1] K.He,X.Zhang,S.Ren,J.Sun, Deep Residual Learning for Image Recognition, IEEE Proc. Computer Vision and Pattern Recognition, pp.770-778, 2016.
- [2] R.T.Q.Chen,Y.Rubanov,J.Bettencourt,D.Duvenau, Neural Ordinary Differential Equations, NIPS'18, pp.6572-6583, 2018.