Deep Learning における液性限界の三次元分布推定における活性化関数の影響

大阪産業大学　学生会員　○梶原　陸
大阪産業大学　正会員　小田和広
鹿児島大学　正会員　伊藤真一

１．はじめに

ボーリング調査では，調査地点の調査すべき深度から土試料を採取し，その試料を使って土質試験を行い，地盤特性を明らかにする．したがって，ボーリング調査結果を三次元的に見れば，空間的にばらついた疎な点の情報の集合体である．この地盤情報を空間的に補間することができれば，未調査点の地盤情報を推定することとなり，それは地盤構造物の設計・施工において非常に有益であろう．既存のボーリング調査結果を使って未調査点の地盤情報を推定しようとする試みの一つに人工知能を用いる方法が国内外で提案されている．人工知能による地盤情報の推定精度は，地盤構成の複雑さ，地盤情報の量および人工知能の表現能力等の関係によって決定される．特に，近年では，表現能力の高い Deep Learning が普及しており，地盤情報の空間補間に対する適用が期待されている．本研究では，Deep Learning を使って大阪市内の沖積粘土層の地盤情報の空間補間を試みる．特に，活性化関数や中間層の数と地盤情報の補間精度の関係について考察する．

２．Deep Learning と活性化関数

Deep Learning とは，最もポピュラー人工知能技術であるニューラルネットワークの一種である．**図-1** は Deep Learning の構造を模式的にしている．ニューラルネットワークは，入力層，中間層，出力層およびそれらを繋ぐニューロンによって構成される．入力層から与えられた入力値は中間層を介すうちに変換され出力層で出力される．中間層では，ニューロンからの入力に対し重み付き総和を行い，その総和値に対して次の層へ伝達する信号の値が決定される．この時，伝達する信号の強さを決定するのが活性化関数である．**図-2** は解析に使用した活性化関数を示している．ReLU 関数は Deep Learning において最もポピュラーに用いられている活性化関数である．その特徴として，入力が正であれば，入力値をそのまま出力として次の層へ伝達する．逆に入力が負であれば，出力値はゼロとなる．一方，Sigmoid 関数は Deep Learning が普及する以前の代表的な活性化関数である．Sigmoid 関数は 0~1 の範囲でしか出力しないことが特徴的である．そのため，現在では，主に二値の判別分析における出力層の活性化関数として用いられる．

３．解析の概要

図-3 は解析に用いたボーリング調査地点を示している．解析対象として，大阪市内の上町台地の西側の地域を選んだ．この地域の沖積粘土は，基本的には完新統における海進によって形成されているが，淀川からの堆

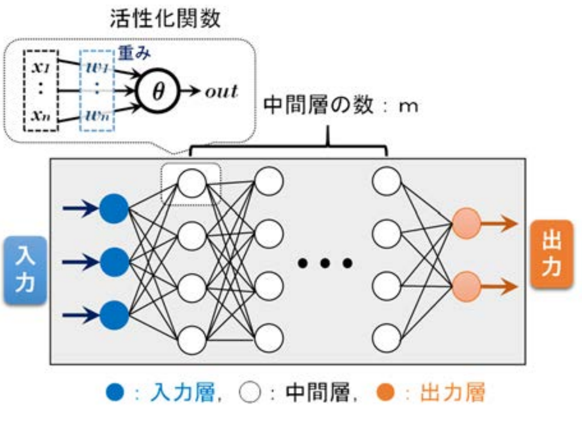

図-1 Deep learning の構造

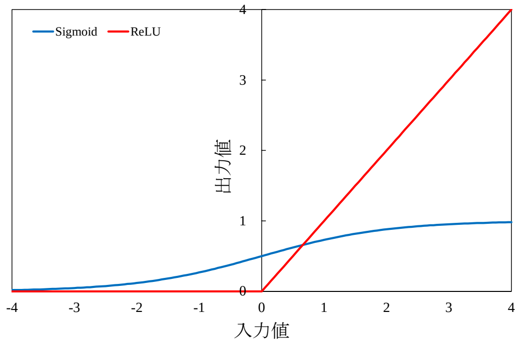

図-2 解析に使用した活性化関数

図-3 解析に用いたボーリング調査地点

キーワード　人工知能，沖積粘土，液性限界，Deep learning，空間補間
連絡先　〒574-8530 大阪府大東市中垣内 3 丁目 1−1　大阪産業大学工学部都市創造工学科　TEL 072-875-3001

© Japan Society of Civil Engineers

表-1　解析条件

活性化関数	ReLU 関数	Sigmoid 関数
ニューロンの数	20	20
中間層の数	1~36	1~5

積物の影響も受けており，大阪湾内の粘土層に比べ堆積構造が複雑である．解析には，学習用（図-3 中の赤丸）として 3161 個，検証用（図-3 中の青丸）として 1356 個のデータを用いた．なお，ボーリング毎に学習用と検証用に分けたのではないため，一本のボーリングにおいても学習用と検証用のデータが混在している．

解析では，入力を緯度，経度および深度の 3 つであり，出力はその地点に対応する液性限界の推定値である．解析において損失関数としては平均二乗誤差，評価関数には平均絶対誤差を選んだ．Epoch 数（繰返し調整回数）は 10000 としたが，全ての解析ケースにおいて 100 回になるまでに損失関数と評価関数の両者ともに一定値に収束した．なお，解析には，誤差逆伝播法バッチ勾配降下法を適用した．**表-1** は Deep Learning の解析条件を示している．活性化関数の種類にかかわらずニューロン数は 20 とした．活性化関数が ReLU 関数の場合，中間層の数を 1~36 まで変化させた．活性化関数が Sigmoid 関数の場合は，中間層の数を 1~5 とした．

４．活性化関数と推定精度

図-4 は検証時の MAE，MARE および RMSE と中間層の数の関係示している．但し，活性化関数として ReLU 関数を用いている．ここで，MAE は平均絶対誤差（Mean Absolute Error），MARE は平均全体相対誤差（Mean Absolute Relative Error）および RMSE は平均平方二乗誤差（Root Mean Square Error）であり，それぞれ式(1)~(3)で与えられる．

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|x_i^P - x_i^O\right| \tag{1}$$

$$MARE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i^P - x_i^O}{x_i^O}\right| \tag{2}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i^P - x_i^O\right)^2} \tag{3}$$

ここに，N：サンプル数，　：推定値，　：実際の値である．中間層の数が 4 であるケースを除けば，中間層の数が 8 までは，MAE，MARE および RMSE のいずれもが単調に減少している．中間層の数が 8 以上ではいずれの指標もわずかではあるが中間層の数が 8 の値よりも大きくなっている．すなわち，中間層が多いほどネットワークの表現能力（データにフィットできる能力）が高くなるため，推定精度は向上すると考えがちであるが，学習時の過学習によって検証時の推定精度が向上することにはならない．**図-5** は検証時の MAE，MARE および RMSE と中間層の数の関係示している．但し，活性化関数として Sigmoid 関数を用いている．MAE，MARE および RMSE は活性化関数に ReLU 関数を用いた場合よりもかなり大きい．また，MARE を除き，中間層の数の影響をほとんど受けない．これは，勾配消失問題と呼ばれるよく知られた現象である．つまり，バッチ勾配効果法を適用する場合，活性化関数の勾配が解析上必要となる．ReLU 関数を用いる場合，入力値が正であれば活性化関数の勾配は非ゼロである．一方，Sigmoid 関数を用いている場合，活性化関数の勾配は入力値がゼロ付近を除きゼロになる．また，活性化関数の勾配の最大値は入力値がゼロの時に 1/4 でしかない．解に対するその影響は中間層の数のべき乗で与えられるので，中間層が多くなってもネットワークの表現能力は高まらない．このことが，一つ前の世代のニューラルネットワークにおいて深層化を妨げていた原因の一つである．

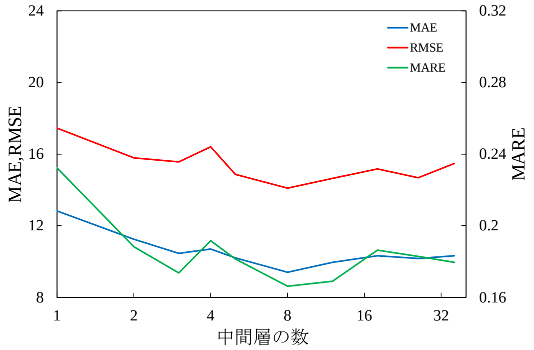

図-4 MAE，MARE および RMSE と中間層の数の関係（検証時，活性化関数は ReLU 関数）

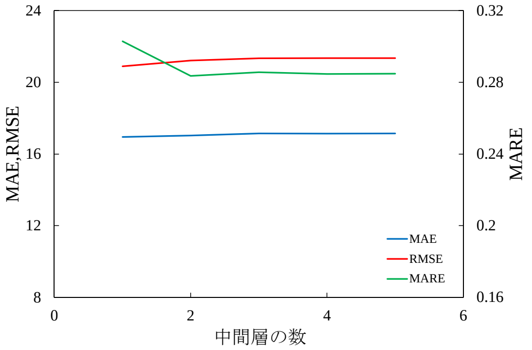

図-5 MAE，MARE および RMSE と中間層の数の関係（検証時，活性化関数は Sigmoid 関数）

５．まとめ

本研究で得られた主な知見は以下の通りである．

1) Deep Learning では中間層の増加は必ずしも検証時の推定精度の向上をもたらさない

2) 活性化関数として Sigmoid 関数を用いた場合，ReLU 関数を用いた場合よりも推定精度は低い．

3) 活性化関数として Sigmoid 関数を用いた場合 Deep Learning はできないことを再確認した．

© Japan Society of Civil Engineers