

論文メモ

文献番号	0012
日付	2021 年 12 月 07 日
名前	武川海斗

文献情報

著者	T. Chai and R. R. Draxler
英文タイトル	Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature
和文タイトル	二乗平均平方根誤差 (RMSE) か平均絶対誤差 (MAE) か？ - RMSE を避けることに対する文献上の議論
書誌情報	Geoscientific Model Development, pp. 1247–1250, 2014
キーワード	なし.

1 論文の要約

RMSE と MAE は誤差情報を基に、モデル評価の指標によく使われる。本論文では、RMSE と MAE の指標の違いについて論じている。先に結論として、どちらが優れているというものではない。RMSE と MAE 複数の測定基準を組み合わせるが必要な場合があります。

2 ベースとなった手法

以下の式は、MAE, RMSE の定義式である。2 式の明確な違いは、絶対誤差か二乗誤差かである。MAE は絶対誤差を扱うのだが、これは最適化の際に問題が起こりやすい。一般に二乗数の方が最適化が容易であり、最適化の面では RMSE の方が有用である。

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |e_i| \quad (1)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \quad (2)$$

3 実験デザイン・結果と考察

RMSE と MAE の性能を比較するために、ランダムに生成したデータを基に実験を行っている。データ数 n を 4, 10, 100, ..., 1000000 まで試した結果を、表としてまとめている。RMSE では、 n が 100 以上になると誤差分布が 5% 以内になるため、「正確な解」に収束している。逆にデータ数が 100 未満の場合、RMSE の誤差が大きく、信頼性が低いと言える。そのため、データ数さえ多ければ信頼性の高い結果を得ることができる。

一方、MAE の場合は、データ数を増やした場合、0.8 に収束している。やはり、MAE の場合もデータ数の量によって信頼性が変わることを本論文では述べている。

しかし、これらのランダムデータは、仮定したガウス分布から生成したものである。RMSE は、二乗誤差を扱うため、極端に外れた値が含まれている場合、過度に反応してしまい、ロバスト性に欠ける。一方、MAE は絶対誤差を扱うため、全てのデータについて平等な影響を受ける。そのため、ロバスト性という面では優れていると言える。しかし、RMSE は MAE と異なり、データごとに「重み」づけすることでモデル性能指標に違いをつけることができる面で、有用であると言える。

これらの議論から、RMSE と MAE のどちらが優れているということはない。そのため、RMSE と MAE を含めたさまざまな評価指標が今後登場することが期待される。