

Recurrent Neural Networkに基づく複数時系列関係を考慮した時系列予測 The time series forecasting using relationship of multi time series based on Recurrent Neural Network

松浦匠吾[†]
Shogo Matsuura

六井淳[‡]
Jun Rokui

1. まえがき

時間構造を持つデータに対して有用な Recurrent Neural Network(以下RNNとする)を用いて時系列データを予測する研究が行われている。松井らはRNNに株式の成長率データをRNNに入力し予測を行う研究を行っている。この研究では訓練データに対して良好な予測精度を示したが、テストデータに対して予測精度を悪化させた[1]。南らはRNNに企業イベントを織り込んだデータを学習させ株価を予測する研究を行った[2]。この研究では、RNNの入力に予測対象のデータのみを入力するのではなく決算発表日やプレスリリース等のイベントを変数に用いることで、株価単体を学習させるよりも精度が向上することが確認された。また、安達らは複数時系列データを利用する方法として、重回帰分析を用いて複数時系列データを集約するフレームワークに関する研究を行い、説明変数の入力数を変化させることでRNNの予測精度の向上を実験的に確認した[3]。この研究は入力する説明変数は1つずつ増加させて検証するものであり、入力する説明変数の組み合わせが課題となっている。

本研究では、グレンジャー因果検定[5]に基づいて重回帰分析で利用する説明変数を抽出し、複数時系列データによって表現されたデータをRNNの入力に用いて予測を行うフレームワークを提案する。

2. 提案手法

提案手法の概略図を図1に示す。

本手法は主に回帰分析部とRNN部に分けられる。回帰分析部では予測対象時系列を目的変数、複数の時系列データから抽出した時系列データを説明変数として重回帰分析を行い、予測対象データの近似時系列を作成する。RNN部では回帰分析部で作成された近似時系列を入力として予測対象時系列データの未来値を予測する。最後に実測値と予測値を用いて予測精度を検証する。はじめにデータ前処理として、それぞれの時系列データに対して、そのデータの最大値で除算することで全てのデータを0から1の範囲にする正規化を行った。

2.1. 回帰分析部

回帰分析部では重回帰分析を行い、複数時系列データを集約して予測対象時系列の近似時系列を作成する。

重回帰分析は k 次元データで $k-1$ 個の変量を説明変数にし、1個の変数を目的変数にして回帰式を求めることができる。予測対象 \hat{Y} は、 $k-1$ 個からなる説明変数

を、 x_1, \dots, x_{k-1} とすると、重回帰モデルは

$$\hat{Y} = \omega_0 + \omega_1 x_1 + \dots + \omega_{k-1} x_{k-1} \quad (1)$$

と表現することができ、 $\omega_1 \dots \omega_{k-1}$ は重回帰モデルにおける説明変数の偏回帰係数である。また、 ω_0 は切片である。このとき実測値を Y 、回帰分析による推定値を \hat{Y} を用いて残差の平方和は ω の関数

$$f(\omega_1 \dots \omega_{k-1}) = \sum_{i=0}^n (Y_i - \hat{Y}_i)^2 \quad (2)$$

で表すことができる。この時残差平方和を最小にする係数 ω を求めることで回帰方程式が求められる[4]。

重回帰分析では説明変数の種類や個数を変化させることで予測精度も変化するが、適切な説明変数を選択することは、データ数が多くなるほど困難になる。そこで、グレンジャー因果検定を用いることで、説明変数の選択を行った。

2.1.1. グレンジャー因果検定

グレンジャー因果検定とは、変数 X, Y 間の因果関係を、過去の X の値が未来の Y の値を予測する際に「有用」であるとき、変数 X を変数 Y の原因と定義するものである[5]。2つの時系列 X, Y において X から Y へグレンジャー因果関係があるか調べるとき、

$$Y(t) = \sum_{i=1}^p a_i Y(t-i) + e_0(t) \quad (3)$$

$$Y(t) = \sum_{i=1}^p a_i Y(t-i) + \sum_{i=0}^p b_i X(t-i) + e_1(t) \quad (4)$$

という2つのモデルを考える。 p は遡るタイムステップであり、 t は時刻、 e はノイズである。式(3)は Y の予測に Y の過去情報のみを用いる自己回帰モデルである。式(4)は Y の過去情報と他の系列 X の過去情報を用いて Y の予測を行うモデルである。この2つのモデルにおいて、前者の予測精度よりも後者の予測精度が向上するとき、 X から Y へグレンジャー因果関係があるという。

本研究では予測対象データに対してグレンジャー因果関係があるデータを説明変数として抽出した。

[†]静岡県立大学経営情報イノベーション研究科, Graduate School of Management and Information of Innovation, University of Shizuoka

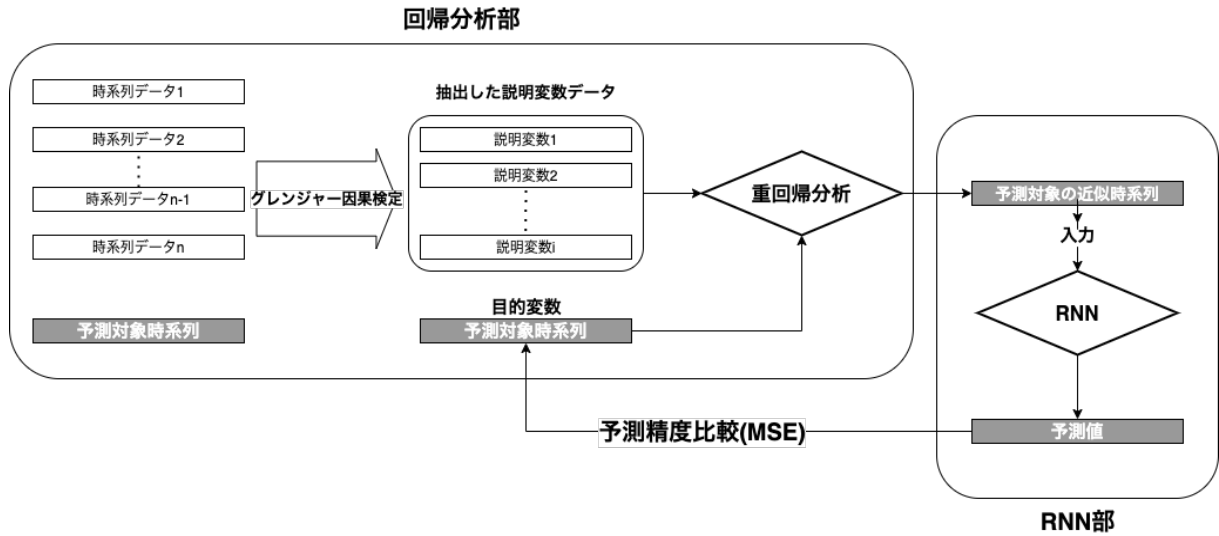


図 1: 提案手法の概略図

2.1.2. 多重共線性

重回帰分析において説明変数内に強い相関関係があると、説明、予測のいずれにおいても支障をきたす多重共線性の課題がある [6]. 多重共線性は、2つ説明変数間の相関係数を r とするとき、ベクトル拡大要因 (Variance Inflation Factor: 以下 VIF とする) を算出し、VIF が 10 よりも大きな値になるとき、多重共線性ありと判断される。VIF は次のように求める。

$$VIF = \frac{1}{1 - r^2} \quad (5)$$

本研究では多重共線性の回避法として、多重共線性ありと判断されたデータの組み合わせの片方を説明変数から除外する。

2.2.RNN 部

ニューラルネットワークの 1 つである Recurrent Neural Network: RNN は構造上の特性として、ネットワークの隠れ層ノードの出力が次の時刻で入力となる再帰的な構造をもつ [7]. この再帰構造によって時系列データの特性を捉えることができる。RNN 部では、回帰分析部で作成された予測対象時系列の近似時系列を入力として予測対象時系列の未来値の予測を行う。本研究では Gated Recurrent Unit と Long Short Term Memory の 2 つの RNN モデルを用いて比較検証を行った。

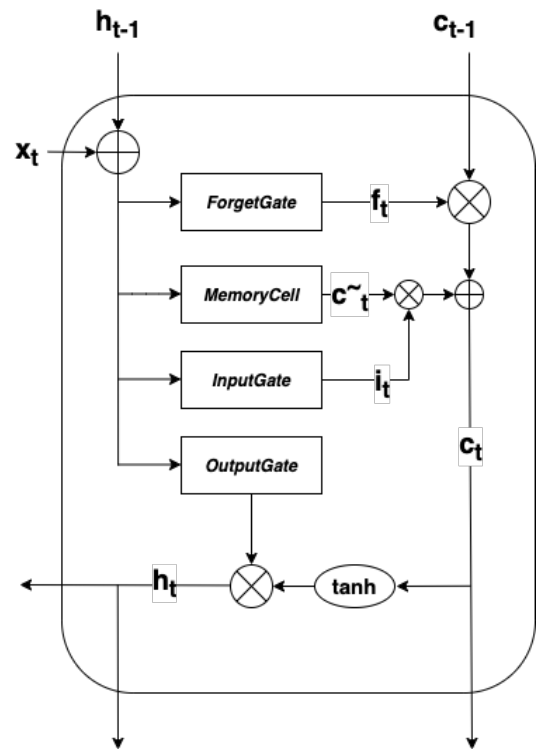


図 2: LSTM セル

2.2.1. Long Short Term Memory

Long Short Term Memory (以下 LSTM とする) は 1997 年に Hochreiter らによって開発された [9]. LSTM の模式図を図 2 に示す。LSTM は忘却ゲート、入力ゲート、出力ゲート、メモリセルから構成されており、各構造の処理を解説する。忘却ゲートではセルの情報の取捨選択の割合を制御する。

入力されるベクトルを x_t 、前時刻 $t-1$ における LSTM の出力ベクトルを h_{t-1} とする。また、各ゲートの重み行列を ω^* 、 u^* 、バイアスベクトルを b^* とすると忘却ゲートベクトル f_t は

$$f_t = \sigma(\omega^f x_t + u^f h_{t-1} + b^f) \quad (6)$$

と表される。次に入力ゲートのベクトルを i_t とすると

$$i_t = \sigma(\omega^i x_t + u^i h_{t-1} + b^i) \quad (7)$$

で表される。入力ゲートではノードに新たに入力される x_t と及び $t-1$ 時刻のノードの出力ベクトル h_{t-1} の情報量を制御する。メモリセルでは情報の記憶を行う。メモリセルの更新をする候補値ベクトルを \tilde{c}_t , 時刻 t におけるメモリセルのベクトルを c_t とすると

$$\tilde{c}_t = \tanh(\omega^c x_t + u^c h_{t-1} + b^c) \quad (8)$$

$$c_t = f_t * c_{t-1} + i_t * \tilde{c}_t \quad (9)$$

で表される。出力ゲートでは LSTM の出力制御を行う。出力ゲートのベクトルを o_t とすると、

$$o_t = \sigma(\omega^o x_t + u^o h_{t-1} + b^o) \quad (10)$$

で表され、時刻 t における出力ベクトル h_t は活性化関数 \tanh で活性化されたメモリセルベクトルと出力ゲートベクトルが掛け合わされ、

$$h_t = o_t \odot \tanh(c_t) \quad (11)$$

で表すことができる。ここで、 \odot はアマダール積であり、行列の要素ごとの積を表す。

2.2.2. Gated Recurrent Unit

Gated Recurrent Unit(以下 GRU とする) は 2014 年に Cho らによって開発された [8]。GRU のモデル図を図 3 に示す。GRU はリセットゲートと更新ゲートの二つのゲート構造で構成される。内部の計算パラメータ数は Long Short Term Memory よりも少なく、計算時間を短縮できることが期待されている。

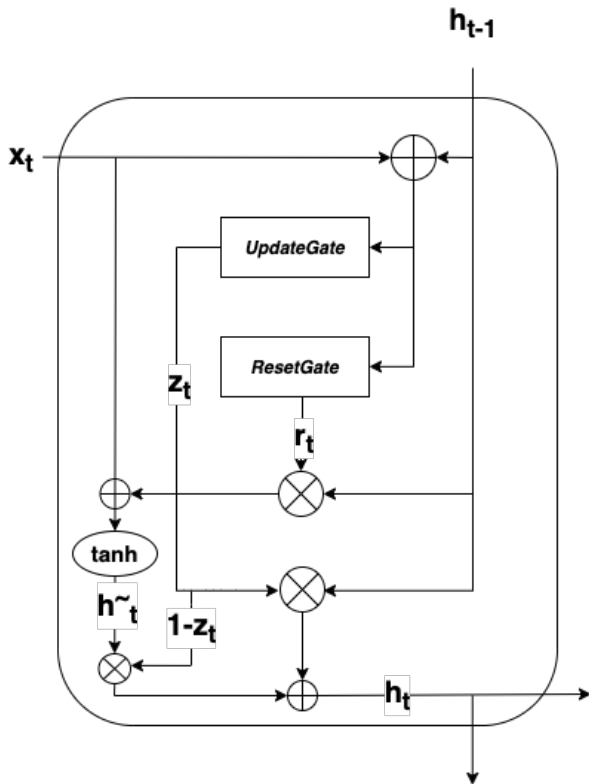


図 3: GRU セル

次に GRU における処理を解説する。各ゲートの重み行列を ω^*, u^* , バイアスベクトルを b^* とする。時刻 t において更新ゲートのベクトル z_t , リセットゲートのベクトル r_t は、入力されるベクトルを x_t , 前時刻 $t-1$ における GRU の出力ベクトルを h_{t-1} とすると、

$$z_t = \sigma(\omega^z x_t + u^z h_{t-1} + b^z) \quad (12)$$

$$r_t = \sigma(\omega^r x_t + u^r h_{t-1} + b^r) \quad (13)$$

と表される。

リセットゲートの値は過去ベクトルと掛け合わされ、入力 x_t とともに活性化関数により活性化される。活性化後のベクトルを \tilde{h}_t とすると、

$$\tilde{h}_t = \tanh(\omega^h x_t + r_t \odot u^h h_{t-1} + b^h) \quad (14)$$

で与えられる。この時、 $\omega^z, u^z, \omega^r, u^r, \omega^h, u^h$ は重み、 b^z, b^r, b^h はバイアスである。GRU ではこの 9 つのパラメータを更新する。最後に、更新ゲートのベクトルは z_t と $(1 - z_t)$ に分割され、それぞれ過去ベクトル h_{t-1} と活性化ベクトル \tilde{h}_t かけられる。時刻 t における出力 h_t は

$$h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \tilde{h}_t \quad (15)$$

と表される。

2.2.3. 精度評価指標

本研究では提案手法及び比較対象の予測精度の評価指標として、平均二乗誤差 (Mean Squared Error: 以下 MSE とする) を用いた。MSE は、RNN が出力した長さ n の予測値 $\hat{Y} = \{\hat{y}_0, \hat{y}_1, \dots, \hat{y}_n\}$ とそれに対応する実測の値 $Y = \{y_1, y_2, \dots, y_n\}$ を用いて

$$MSE = \frac{1}{n} \sum_{i=0}^n (y_i - \hat{y}_i)^2 \quad (16)$$

で計算される。

3. 検証実験

本研究の提案手法の回帰分析部において、多重共線性を考慮に入れた場合と考慮に入れない場合、及び RNN に自己系列のみを入力し予測する 3 種類の手法による MSE 値の比較検証を行う。また、RNN モデルに GRU と LSTM のそれぞれを用いた場合についても比較検証を行う。実験に使用するデータは日経平均株価や米ドル円レートなど、1988 年 1 月から 2005 年 3 月までの 1 ヶ月ごとのデータ 26 種類を用いる。データ長は 207 である。データは日本銀行 [10]、内閣府 [11]、セントルイス連邦準備銀行 [14] 及び e-Stat [12] から入手した。本研究で用いたデータの略称、正式名称、及び出典は図 4 の通りである。RNN モデルに GRU と LSTM を用いてそれぞれの手法で 10 回ずつ実験を行い、最も優れた MSE の値をそのデータの予測精度とする。

次に実験を行う 3 種類の手法について解説する。

略称	時系列データの名称	出典
cgpi	企業物価指数	日本銀行
ci	コンポジットインデックス	内閣府
di	ディフュージョンインデックス	内閣府
cpi	消費者物価指数	e-Stat
japanese_number	総人口	e-Stat
kokogyo	鉱工業生産付加価値指数	経産省
M0	機械受注額	内閣府
mbase	マネタリーベース	日本銀行
nikkei	日経平均株価(終値月平均)	セントルイス連邦準備銀行(FRED)
usd_jpy	米ドル/円レート	日本銀行
situgyou	失業率	e-Stat
prole	労働人口	e-Stat
sangyou	全産業活動指数	経産省
tingin	実質賃金	e-Stat
koyou_s	常時雇用指数	e-Stat
work_time	実労働時間	e-Stat
export_index	輸出物価指数	日本銀行
import_index	輸入物価指数	日本銀行
M1	マネーストック統計M1	セントルイス連邦準備銀行(FRED)
M2	マネーストック統計M2	セントルイス連邦準備銀行(FRED)
M3	マネーストック統計M3	セントルイス連邦準備銀行(FRED)
cpi_usa	消費者物価指数(米国)	セントルイス連邦準備銀行(FRED)
M2_usa	米国マネーストック統計M2	セントルイス連邦準備銀行(FRED)
situgyou_usa	失業率(米国)	セントルイス連邦準備銀行(FRED)
kpremium_num	雇用保険基本手当受給資格決定数	e-Stat
kyuzin	有効求人倍率	e-Stat

図 4: 実験データの名称と出典

3.1. 自己系列入力手法

予測対象時系列の過去情報を RNN に入力し, 予測を行う従来の手法を自己系列入力手法と題する. システムは正規化された予測対象時系列を学習データ: 教師データ = 8 : 2 に分割し, 学習データを RNN に入力して予測する. 予測値と教師データから MSE 値を算出し出力する.

3.2. グレンジャー適応手法

提案手法である重回帰分析による近似時系列を用いた RNN 予測手法について解説する. RNN の入力に重回帰分を行うことで得られた予測対象時系列の近似時系列を用いる. グレンジャー因果検定に基づいて重回帰分析で扱う説明変数を抽出する. この手法をグレンジャー適応手法と題する. 図 5 に処理の流れを示す.

グレンジャー因果検定を適応した結果, それぞれのデータに最適な説明変数を抽出することができた. 図 6 に各データに対応する説明変数の個数を示す.

3.3. 多重共線性回避手法

提案手法において, 多重共線性を考慮した場合においても検証を行う. 図 7 に多重共線性回避手法における処理のフローチャートを示す. 多重共線性が発生する説明変数の組み合わせがある時, 片方のデータを説明変数から除外した新たなデータセットを作成することでこの問題を回避する. この手法を多重共線性回避手法と題する. 本研究で使用した時系列データの中には, 2 組 4 種類のデータで多重共線性を発生する強い相関を持つデータが観測された.

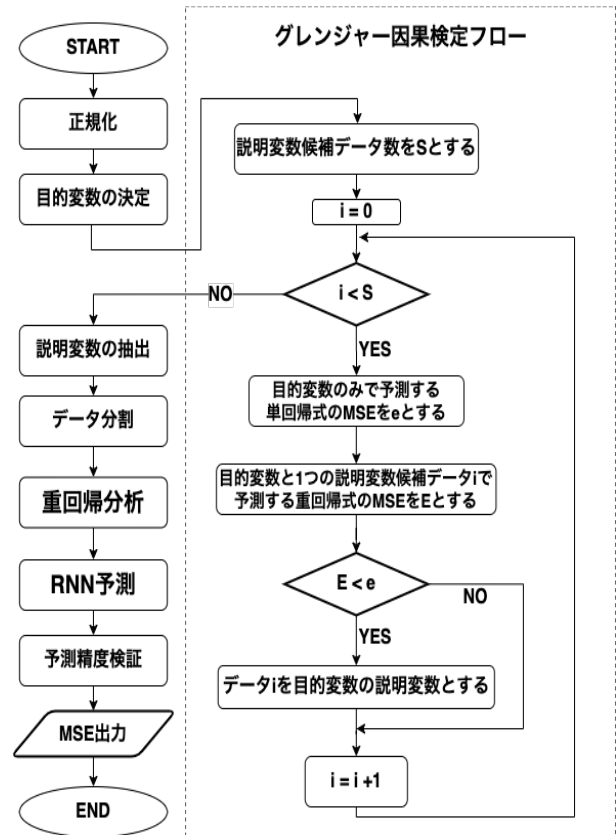


図 5: グレンジャー適応手法の処理の流れ

名称	抽出個数	名称	抽出個数
cgpi	9	tingin	4
ci	12	koyou_s	11
di	9	work_time	12
cpi	10	export_index	5
japanese_number	8	import_index	5
kokogyo	7	M1	4
M0	9	M2	5
mbase	8	M3	8
nikkei	10	cpi_usa	5
usd_jpy	5	M2_usa	13
situgyou	9	situgyou_usa	6
prole	17	kpremium_num	7
sangyou	7	kyuzin	18

図 6: 説明変数の抽出個数

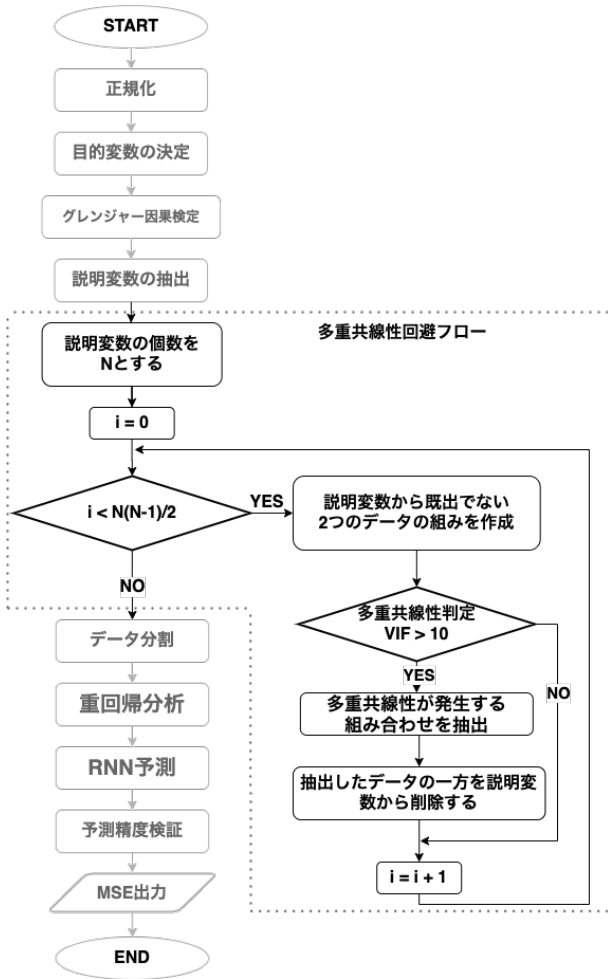


図 7: 多重共線性回避手法の処理の流れ

3.4.GRU による予測実験

RNN モデルに GRU を用いて実験を行った結果を図 8 に示す。図中の緑欄は他の比較している 3 手法の中で最も予測誤差が小さいことを表している。26 種類中 11 の時系列データにおいて本手法である多重共線性回避手法の予測誤差が小さく、フレームワークが有用な結果を示すデータが確認された。

グレンジャー適応手法に関しては良好な MSE を示した結果はなかったが、グレンジャー因果関係に基づいて抽出したデータから更に多重共線性を排除することで、26 種類中 25 の時系列データで予測精度の向上が確認された。

3.5.LSTM による予測実験

RNN モデルに LSTM を用いて同様の実験を行った結果を図 9 に示す。LSTM を用いた実験では 26 種類中 13 の時系列データで提案手法による予測精度が良好な値を示した。今回自己系列入力手法が最も良い精度となった 13 種類のデータは、GRU を用いた時でも同様に自己系列入力手法で最も良好な値を示している。本手法の適応は全てのデータに有効ではなく、データの中に

GRU	自己系列入力	グレンジャー適応	多重共線性回避
cgpi	0.0023932121	0.010670078	0.004299461
ci	0.0039558290	0.332298085	0.001808642
di	0.1146960685	2.37137711	0.134778986
cpi	0.0001807547	0.007218934	0.011547278
japanese_number	0.0003987605	0.000202754	0.00018777
kokogyo	0.0037591231	0.525525833	0.019920064
MO	0.0102597678	0.009829236	0.008696616
mbase	0.5901165498	0.075690788	0.022145572
nikkei	0.0388227130	0.835068522	0.201904855
usd_jpy	0.0046461215	0.375132412	0.036410052
situgyou	0.0124948804	0.022533723	0.005161639
prole	0.0001642706	0.009126145	7.02204E-05
sangyou	0.0014513493	0.001866105	0.000287056
tingin	0.0670908449	6.410414782	0.045007053
koyou_s	0.0009737515	0.018490801	0.001073537
work_time	0.0003601529	0.011108795	0.00068484
export_index	0.0153888914	0.022172335	0.021561537
import_index	0.0040321445	0.906245315	0.02656943
M1	0.0715095015	0.357377441	0.018414917
M2	0.0020856270	0.005815557	0.000239522
M3	0.0001765832	0.012054319	0.001367988
cpi_usa	0.0002999311	0.085709791	0.002313205
M2_usa	0.0031412212	0.165744982	0.006462189
situgyou_usa	0.0587082018	3.139331653	0.001593844
kpremium_num	0.0172183447	0.108801067	0.018468296
kyuzin	0.0329103411	0.037283899	0.07484221

図 8: GRU を用いた各手法の MSE 値

は重回帰分析の適応が RNN の予測精度に悪影響を及ぼすものが確認された。

3.6.GRU – LSTM 比較

グレンジャー適応手法において、GRUの方が予測精度が良好だった時系列データは 26 種類中 10 種類であった。これに対して、多重共線性回避手法では、GRU が LSTM を上回った時系列データは 26 種類中 19 種類と大幅に増加していることから、GRU を用いた提案フレームワークでは、多重共線性回避手法が予測精度向上に影響していると考えられる。

4. おわりに

本研究ではリカレントニューラルネットワークを用いた時系列データ予測において、複数時系列データを重回帰分析と組み合わせることで集約し入力するフレームワーク、及び本フレームワークにおける説明変数選択の効率化手法としてグレンジャー因果検定に基づく抽出を行う仕組みを提案した。検証実験では RNN モデルに GRU と LSTM を利用してモデル間の予測精度、計算速度について比較を行った。GRU を用いた実験では多重共線性回避を適応することで、提案手法における予測精度を向上させることが確認された。しかし、自己系列入力手法が最も良好な値を示した時系列データが 2 つの RNN モデルで共通していることから、本手法に適さないデータも存在することも明らかとなった。本研究

LSTM	自己系列入力	グレンジャー適応	多重共線性回避
cgpi	0.0039643213	0.011201426	0.0068754
ci	0.0033136238	0.332735905	0.002033662
di	0.1130809709	2.254469499	0.137669295
cpi	0.0000328241	0.003488836	0.014215736
japanese_number	0.0002653977	0.000206733	0.015410347
kokogyo	0.0036365162	0.117683909	0.01986006
MO	0.0102412167	0.003363313	0.00879419
mbase	0.4725116268	0.092791819	0.022246584
nikkei	0.1144580055	0.982358505	0.181644412
usd_jpy	0.0079170300	0.398629543	0.026112391
situgyou	0.0051732635	0.022964806	0.005096271
prole	0.0000257801	0.002559307	0.013829164
sangyou	0.0009361345	0.00085123	0.004422483
tingin	0.0681604276	7.203912359	0.047294885
koyou_s	0.0009468497	0.018591685	0.001142617
work_time	0.0005734994	0.001879567	0.001098721
export_index	0.0127887524	0.021970999	0.02341278
import_index	0.0008557862	0.796082059	0.039699692
M1	0.0786028200	0.354035053	0.017240129
M2	0.0131897223	0.004537225	0.000241036
M3	0.0006027081	0.000416278	0.000234601
cpi_usa	0.0105467880	0.078070364	0.004387513
M2_usa	0.0118588594	0.165348211	0.007704405
situgyou_usa	0.0442836159	2.784164601	0.001960456
kppremium_num	0.0132456684	0.042140779	0.015931713
kyuzin	0.0182353139	0.136768905	0.107855201

図 9: LSTM を用いた各手法の MSE 値

で行った多重共線性回避の方法は一方の時系列データを削除する次元削減を用いた。これはデータ数が増えると汎用性を失うため、説明変数データの組み合わせ最適化手法の検討が今後の課題として挙げられる。

参考文献

- [1] 松井藤五郎, and 汐月智哉. "LSTM を用いた株価変動予測." 人工知能学会全国大会 論文集 2017 年度人工知能学会全国大会 (第 31 回) 論文集 2D3-OS-19a-2, pp.1-2, 一般社団法人 人工知能学会, 2017.
- [2] 南正太郎. "LSTM-RNN を用いたイベント考慮後の株価時系列予測." 人工知能学会 金融情報学研究会 (第 19 回) 資料, pp.42-44. 一般社団法人 人工知能学会, 2017.
- [3] 安達 凜, 六井 淳. "複数入力を用いた Recurrent Neural Network に基づく時系列予測" FIT2019 講演論文集, 2, pp.45-52, 2019
- [4] 石井俊全, 統計学大百科事典, 初版, 日経印刷, 2020, pp.188-189
- [5] 近原 鷹一, "時系列データからの因果関係の発見", NTT コミュニケーション科学基礎研究所, 2019
- [6] 大垣俊一, 重回帰と偏相関 その後, Argonauta, 18, pp.17 - 30, 2010.

[7] 麻生川稔. "時系列信号を認識するリカレント・ニューラルネットワーク." 全国大会講演論文集 人工知能および認知科学, pp.410-411, 1989.

[8] Kyunghyun Cho, Bart van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, Yoshua Bengio, Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation, EMNLP, 2014

[9] Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, Neural Computation, Vol. 9, No. 8, pp. 1735 - 1780 (1997)

[10] 日本銀行. "日本銀行時系列統計データ検索サイト". <https://www.stat-search.boj.or.jp/>.

[11] 内閣府. "内閣府 統計情報・調査結果". <https://www.esri.cao.go.jp/index.html>.

[12] 総務省. "e-Stat 政府統計の総合窓口". <https://www.e-stat.go.jp/>.

[13] 経済産業省. "経済産業省統計". <https://www.meti.go.jp/statistics/index.html>.

[14] Federal Reserve Bank of St. Louis. Federal Reserve Economic Data — FRED — St. Louis Fed. <https://fred.stlouisfed.org/>.