

1. はじめに
2. 金融経済の要因と因果関係
3. LSTM による時系列予測
3. おわりに

金融経済要因の波及効果を考慮した 投資ストラテジー構築による 自動売買への適用

戸田 真聰

富山県立大学 情報システム工学科

2024年10月8日

1. はじめに

2/13

背景

1. はじめに
2. 金融経済の要因と因果関係
3. LSTM による時系列予測
3. おわりに

1996 年の外国為替 (Foreign Exchange: FX) 取引の完全自由化により FX 取引が誕生してから、年々金融市场の規模は拡大している。現在では多くのトレーダーが、コンピュータを駆使することで自動的にルールに従い取引をするシステムトレードを導入している。また昨今は人工知能を導入することで価格の予測、戦略を獲得するという研究や、金融市场への他市場からの影響を調べている研究は存在するものの、為替市場においてそれらを用いて自動売買をする研究は少ないよう見受けられる。

目的

本研究では、リアルタイムで取得した Tick データを使用して、 VAR-LiNGAM による時系列を考慮した因果探索を行い、円ドル為替市場に対して他市場が因果性を持つのかを確認し、その結果を考慮して LSTM による機械学習を行うことで未来の値動きを予測し、自動売買を行い利益を出すことを考える。

2. 金融経済の要因と因果関係

3/13

VAR-LiNGAM

1. はじめに
2. 金融経済の要因と因果関係
3. LSTM による時系列予測
3. おわりに

MT5 から Tick データを取得し、正規化などデータの前処理を行い、
VAR-LiNGAM による時系列を考慮した因果探索を行う。

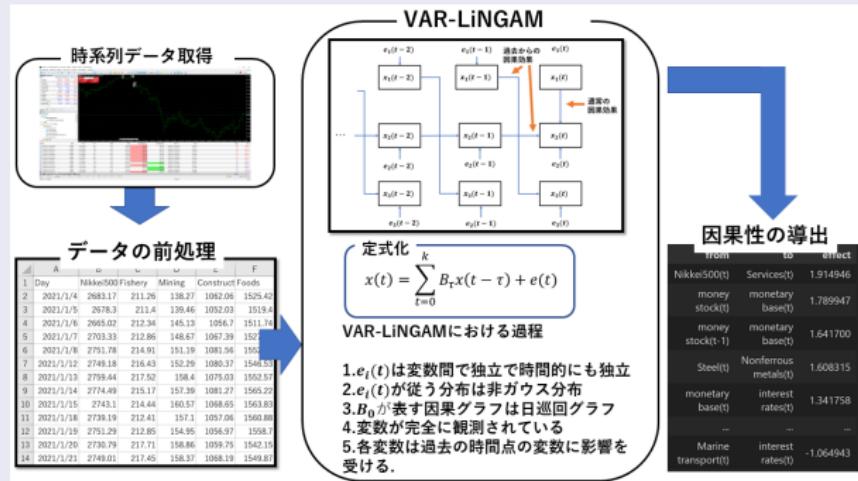


図 1: システムの流れ

VAR-LiNGAM

円ドル為替市場に対して、金、日経平均株価、プラチナなどが因果性を持つのかを確認する

- はじめに
 - 金融経済の要因と因果関係
 - LSTMによる時系列予測
 - おわりに

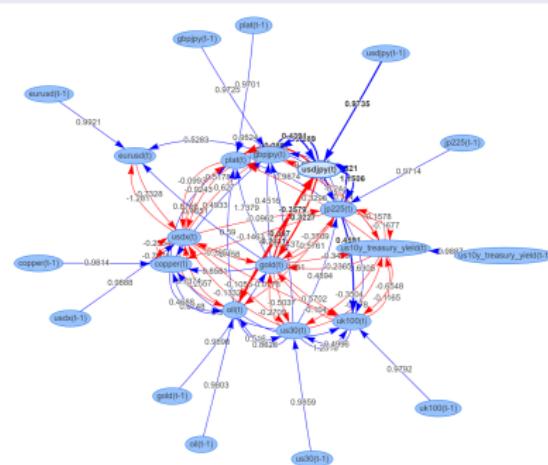


図 2: VAR-LINGAM の実行結果

3. LSTMによる時系列予測

5/13

LSTM

- RNN(Recurrent Neural Network)とは、ニューラルネットワークの一種であり、特に時系列データやシーケンスデータの解析に特化したモデル。
- LSTM(Long Short-Term Memory)とは、RNN の一種であり、過去の情報を長期記憶しておく記憶セルの導入により、RNN が持っていた「長期記憶の消失」というデメリットをある程度改善したもの。

1. はじめに
2. 金融経済の要因と因果関係
3. LSTM による時系列予測
3. おわりに

3. LSTMによる時系列予測

6/13

- はじめに
- 金融経済の要因と因果関係
- LSTMによる時系列予測
- おわりに

交差検証

交差検証 (cross-validation) とは、データセットを複数に分割し、「訓練データ」による予測モデルの学習と「検証データ」による評価を、複数回実施する方法。

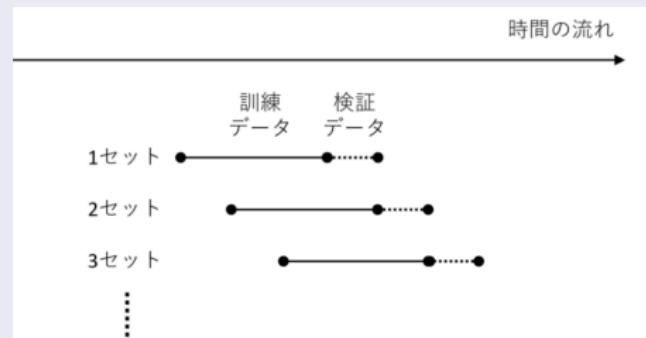


図 3: 交差検証のイメージ

3. LSTMによる時系列予測

7/13

データの取得

1. はじめに
2. 金融経済の要因と因果関係
3. LSTMによる時系列予測
3. おわりに

MT5からドル円, 金, 日経平均株価, プラチナなど11個の項目の, 2014年9月25日から2024年9月25日までの日足データを取得了.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	<DATE>	usd/jpy	gold	jp225	us30	oil	copper	uk100	gbp/jpy	eur/usd	plat	usdx	Unnamed: us10y_treasury_yield		
2	2014/9/25	108.736	1221.09	16109	16998.5	92.44		6637.7	177.408	1.27496		85.3	0	2.511	
3	2014/9/26	109.267	1217.67	16352	17111.5	93.38		6669.4	177.489	1.26807		85.75	1	2.535	
4	2014/9/28			16365	17096.5			6663.8							
5	2014/9/29	109.485	1215.47	16239	17057.5	94.28		6638.1	177.802	1.26842		85.72	2	2.491	
6	2014/9/30	109.619	1207.92	16166	17022	91.3		6601.1	177.709	1.26295		86	3	2.508	
7	2014/10/1	108.878	1213.48	15872	16823	90.67		6541.2	176.215	1.26209		85.94	4	2.403	
8	2014/10/2	108.406	1214.24	15555	16795.5	91.36		6472.2	175.009	1.26673		85.71	5	2.438	
9	2014/10/3	109.773	1191.47	15900	17004.5	89.69		6528.3	175.284	1.25146		86.76	6	2.447	
10	2014/10/5			15927	17028.5			6539.3							
11	2014/10/6	108.77	1206.89	15765	16964	90.43		6559.3	174.93	1.26538		85.83	7	2.425	
12	2014/10/7	108.015	1208.63	15507	16709.1	88.41		6460.5	173.838	1.26676		85.72	8	2.35	
13	2014/10/8	108.07	1220.83	15688	16975.5	87.7		6537.2	174.715	1.27321		85.36	9	2.33	

図 4: 取得したデータの一部

3. LSTMによる時系列予測

8/13

データの前処理

- はじめに
- 金融経済の要因と因果関係
- LSTMによる時系列予測
- おわりに

- 目的変数：ドル円の終値
- 特徴量：金, 日経平均株価, ダウジョーンズ, WTI 原油, ハイグレード銅, 英 FTSE, 英ポンド/日本円, ヨーロ/米ドル, プラチナ, 米ドル指数, 米国 10 年国債利回りの 12 個の項目の終値
- データの分割：全データを 3 分割して、それぞれで 7 割を訓練データ, 3 割を検証データとする
- 12 個の特徴量のうち、一つでも欠損値がある日は削除する
 - データの正規化：データを 0 から 1 の範囲にする最小-最大スケーリングという方法を用いる。

3. LSTMによる時系列予測

9/13

- はじめに
- 金融経済の要因と因果関係
- LSTMによる時系列予測
- おわりに

モデルの構造

- LSTM 層：3 つの LSTM 層から構築されている。各 LSTM 層は 32 ユニットを持つ。
- ドロップアウト層：各 LSTM 層の後に 2 層ある。ドロップアウト率は 0.2 に設定している。
- 全結合層：ユニット数は 1 に設定している。
- モデルのコンパイル：最適化アルゴリズムを Adam に設定し、学習率は 0.0005 である。
- モデルの学習：今回、エポック数を 150、ミニバッチサイズを 64 に設定している。

予測

今回は、過去 60 日間のデータを使用して、翌日のドル円の終値が前日の終値よりも上がっているか(下がっているか)を予測する

3. LSTMによる時系列予測

10/13

モデルの評価

正解率 (Accuracy) : 全予測に対する正答率

$$\frac{TP + TN}{TP + FP + FN + TN}$$

	実際は正 (Positive)	実際は負 (Negative)
予測が正 (Positive)	TP(真陽性) True Positive	FP(偽陽性) False Positive 第1種の誤り
予測が負 (Negative)	FN(偽陰性) False Negative 第2種の誤り	TN(真陰性) True Negative

図 5: 混合行列

3. LSTMによる時系列予測

11/13

モデルの評価

- はじめに
- 金融経済の要因と因果関係
- LSTMによる時系列予測
- おわりに

適合率 (Precision)：正と判断したデータのうち、実際に正であるものの割合

$$\frac{TP}{TP + FP}$$

再現率 (Recall)：実際に正であるもののうち、正であると予測されたものの割合

$$\frac{TP}{TP + FN}$$

F 値 (F-measure)：適合率と再現率の調和平均

$$\frac{2 * Precision * Recall}{Precision + Recall}$$

3. LSTMによる時系列予測

12/13

実行結果

- はじめに
- 金融経済の要因と因果関係
- LSTMによる時系列予測
- おわりに

全フォールドにおける平均正解率: 0.53
全フォールドにおける平均適合率: 0.51
全フォールドにおける平均再現率: 0.51
全フォールドにおける平均F1スコア: 0.51

図 6: 特徴量がドル円のみの結果

全フォールドにおける平均正解率: 0.54
全フォールドにおける平均適合率: 0.54
全フォールドにおける平均再現率: 0.54
全フォールドにおける平均F1スコア: 0.54

図 7: すべての特徴量を含んだ実行結果

全フォールドにおける平均正解率: 0.55
全フォールドにおける平均適合率: 0.54
全フォールドにおける平均再現率: 0.54
全フォールドにおける平均F1スコア: 0.54

図 8: 因果性が強かった4つ特徴量(金, 日経平均株価, WTI原油, 英ポンド/日本円)を含んだ実行結果

3. おわりに

13/13

今後

- ・ポスターの作成
- ・予測精度をあげる(テクニカル指標などの特徴量を追加, 機械学習モデルの理解・改善)
- ・リアルタイムデータでの適用

1. はじめに
2. 金融経済の要因と因果関係
3. LSTM による時系列予測
3. おわりに