

卒業論文

市場間分析を活用した 高頻度データに対するパラメータ選択による 最適なストラテジー構築

Optimal strategy construction
by parameter selection
for high frequency data using inter-market analysis

富山県立大学 工学部 電子・情報工学科

1815031 木下 大輔

指導教員 奥原 浩之 教授

提出年月: 令和4年 2月

目次

図一覧	ii
表一覧	iii
記号一覧	iv
第1章 はじめに	1
§ 1.1 本研究の背景	1
§ 1.2 本研究の目的	2
§ 1.3 本論文の概要	2
第2章 高頻度データ収集	4
§ 2.1 取引プラットフォーム	4
§ 2.2 インジケータを用いたテクニカル分析	4
§ 2.3 ストラテジーテスターにおけるバックテストと最適化	6
第3章 インジケータの選択	8
§ 3.1 曜日時間帯による変動の特徴の抽出	8
§ 3.2 市場データの選択と活用	8
§ 3.3 因果関係分析による他市場の活用	10
第4章 提案手法	12
§ 4.1 機械学習のためのデータセット作成	12
§ 4.2 機械学習によるルール作成	14
§ 4.3 提案手法のアルゴリズム	14
第5章 数値実験並びに考察	17
§ 5.1 数値実験の概要	17
§ 5.2 実験結果と考察	19
第6章 おわりに	20
謝辞	21
参考文献	22

図一覽

3.1	月曜日	9
3.2	火曜日	9
3.3	水曜日	9
3.4	木曜日	9
3.5	グレンジャー因果性検定	11
4.1	提案手法の流れ	15

表一覽

5.1 実験結果	19
--------------------	----

記号一覧

以下に本論文において用いられる用語と記号の対応表を示す.

はじめに

§ 1.1 本研究の背景

1996年の外国為替証拠金取引（Foreign Exchanger: FX）の完全自由化によりFX取引が誕生してから、年々金融市場の規模は拡大している。通信情報技術の発達と金融工学の進歩は、取引単位の小口化と取引手数料の低下により金融市場への参加者を増やし、取引の簡易化と高速化により金融市場全体の流動性を高めた。この2点は本来の資産クラスを超えた取引も容易にした。このことは外国為替市場においてもさらなる流動性と市場への参加者をもたらし、元々巨大であった外国為替市場はより一層巨大な市場へと変貌した [1]。

通信情報技術の発達がもたらしたの外国為替市場の規模拡大だけでなくトレーダーにも変化をもたらししている。コンピュータが誕生する前や、今ほど性能がない時代はトレーダーの経験や勘といった自身の判断で取引を行う裁量トレードといった取引手法が主であった。コンピュータの性能向上によりそれらを駆使することで自動的にルールに従いトレードを行うといったようなシステムトレードといった物も行われるようになった。また昨今では人口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もおこなわれている [2]。

市場の予測を身近で行う例として投資があげられる。従来の投資の判断基準として用いられているのが金融市場の要因のみによって得られた分析結果である。そこで用いられる分析は、過去の市場の動きから指標を算出して未来の市場の動向を予測するようなものである。

予測に用いられる指標は、直近の市場のデータから一定の値を導出し、その値をもとに今後の市場の値動きを予測することを目的としている。また、指標を単独でもちいて予測を行うことももちろんできるが、それらを複数組み合わせることでより正確な予測を行うことも進められている。また投資を行う上で、このような金融市場のメカニズムを利用して判断することが一般的である。従来の投資の判断基準として用いられているのが金融市場の要因のみによって得られた分析結果である。そこで用いられる分析は、過去の市場の動きから指標を算出して未来の市場の動向を予測するようなものである。

市場内の要因を分析するだけでなく他市場からの影響と市場の変動との関連を分析する研究も出てきているように様々な手法を用いて市場の予測を行おうとする研究が行われている。各企業が出している四季報や政治情勢などの市場外的要因も考慮して予測を行う研究もある。市場の変動要因は何であるかについて検証することで市場の変動要因間の相互関係を明らかにするといったような研究も行われている [3]。

§ 1.2 本研究の目的

従来の投資の判断基準として用いられているのが金融市場の要因によって得られた分析結果である。そこで用いられる分析は、過去の市場の動きから指標を算出して未来の市場の動向を予測するようなものであり、現在の市場の動きに合わないことがある。また為替市場が他の市場に与える影響を分析する研究や、他の市場が為替市場にどのような影響を与えるか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の情報を考慮して売買をする研究が少ないように見受けられる。

そのため本研究では、市場に対しての分析を行うとともに、他市場が為替市場に与える影響を取引の際に考慮することで市場内の要因から得られる分析結果のみならず、そこからでは予測できないような変動にも対応ができるような手法を提案する。

この目的を達成するために、市場を分析するとともに、他市場が為替市場に与える影響を調査し、自動売買するにあたりその影響も考慮するために必要な仕組みを考える。本研究では、時系列解析における代表的なモデルであるベクトル自己回帰モデル (Vector Autoregressive Model: VAR) において、複数間の時系列間の因果関係を同定するグレンジャー因果性分析を用いた為替の自動売買のための分析手法の提案を行う。

まず、為替の取引プラットフォームからリアルタイムの高頻度データを取得し、そのデータを蓄積することで分析に使うヒストリカルデータを作成する。その後、そのヒストリカルデータにおいて分析に用いるインジケータの評価値が最も高くなるようなパラメータを抽出することで、市場に対する売り買いの判断を行う。他市場に対しても同様の分析を行うことでその市場に対する売り買いの判断を行う。並行して他市場から為替市場に対する影響についても分析することで、市場内の要因と市場外の影響を考慮した市場予測システムを作成する。

提案手法によって構築されたシステムの有効性を検証する必要がある。そのために、実際に以上のことを考慮した分析手法を基に自動売買システムを構築する。そして、構築した自動売買システムによって為替の売買を行い、収益を出すことで本研究の提案手法の有意性を示す。また、本研究では作成した自動売買システムをもちいて、実際にデモ口座を使ったりリアルタイム取引を行う。

§ 1.3 本論文の概要

本論文は次のように構成される。

第1章 本研究の背景と目的について説明する。背景では金融市場と通信情報の発達による規模の変化やトレードの仕方の変化を述べた後に、市場の予測を行うために市場内の要因を用いるする方法と他市場が為替市場に与える影響について述べた。目的は市場の動向を予測するためにそれらの情報を考慮した手法を提案し、どのように有効性を示すかを述べた。

第2章 為替取引に使用される取引プラットフォームや用語の説明についてまとめる。また、市場の予測に使われる分析手法の例とその最適化に使われるシステムについて述べる。

第3章 市場における時間帯による特徴の変化と為替市場にどのような市場が影響を与えるのかを述べる．また，どのようにそれらを考慮するのかについて述べる．

第4章 提案手法中の Tick データからヒストリカルデータを作成する部分と，その後，本研究の提案手法の流れについて述べる

第5章 提案手法に基づいて自動売買システムを構築して，運用テストを行う．そして，本研究の提案手法によって得られた結果が有意であることを示す．

第6章 本研究で述べている提案手法をまとめて説明する．また，今後の課題について述べる．

高頻度データ収集

§ 2.1 取引プラットフォーム

取引プラットフォームとは、FX取引をおこなうためのツールを指す。チャートを表示してテクニカル分析やシステムツールを使用したり、PCやスマートフォン端末のアプリから売買の注文を行ったり、約定履歴や保有ポジションを確認したりするのに使用される。

今回の実験において用いる取引プラットフォームはMetaTrader 5(MT5)である。MT5はFXにおけるテクニカル分析及び取引業務を行うトレーダ向けの無料アプリケーションである。MT5は現在世界で最も利用されている取引プラットフォームであり様々な機能が存在する。

MT5はデモ口座を開き非常にリアルな取引を行うことができる。Pythonを使用することでMT5からTickデータの取得や取引のオーダーを送ることで実際にMT5で取引シミュレーションを行うことができる。

§ 2.2 インジケーターを用いたテクニカル分析

インジケーターは過去の価格や出来高などの要素から未来の価格を予測する分析手法である。短期トレードで利益を狙うデイトレード、スキャルピングといった手法に用いられることが多い。テクニカル分析の分析対象は、市場内要因や銘柄別要因であることから過去のデータを用いて分析することが多い。

テクニカル分析で用いられる情報としては、「現在の相場のトレンド傾向・強さ」、「トレンドの転換点」、最近では、時系列データの予測が得意であるニューラルネットワークや機械学習によって大規模なデータから分析を行ったり、最適なテクニカル指標を算出して効果的な予測を行うような研究が行われている。

インジケーターとは、為替レートの時系列情報を計算して売買の判定に利用する指標のことを指す。インジケーターを使用することにより、人間が見るだけではわからない情報が発見できる時がある。インジケーターにはオシレーター系とトレンド系の二つある。為替レートは上がり過ぎると下がる、下がり過ぎると上がる性質がある。オシレーター系のインジケーターはこの性質を活かし、為替レートが上がり過ぎと下がり過ぎを数値化、グラフ化するものである。トレンド系のインジケーターは現在の相場が上がりやすい傾向であるか下がりやすい傾向であるかグラフ化し、視覚的にわかりやすくするものである。

指数移動平滑平均 (Exponential Moving Average: EMA)

一般的な移動平均線は、単純にある期間における平均から算出したものであるが、EMAでは直前の為替のレートを重視するように工夫した指標である。期間 n の単純移動平均 (Simple Moving Average: SMA) は p_t を時刻 t の終値とすると、

$$SMA = \frac{p_t + p_{t-1} + p_{t-2} + \cdots + p_{t-(n-1)}}{n} \quad (2.1)$$

となる。一方、期間 n の EMA は 1 日目は SMA と計算方法は同じだが 2 日以降は、

$$EMA_t = EMA_{t-1} + \alpha(p_t - EMA_{t-1}) \quad (2.2)$$

ここで、

$$\alpha = \frac{2}{n+1} \quad (2.3)$$

そして、EMA で取引を行う際には短期移動線と長期移動線の 2 つを用いる。基本的には、以下のゴールデンクロスとデッドクロスを指標としてルールを作成する。

ゴールデンクロス

短期移動線が長期移動線を下から上へ突き抜けたときに買いサインとなる。

デッドクロス

短期移動線が長期移動線を上から下へ突き抜けたときに売りサインとなる。

ボリンジャーバンド

ボリンジャーバンドは、「標準偏差」と「正規分布」に基づいた考え方であり、68-95-99 ルールを使って取引のルールを決める。これは、正規分布は ± 1 標準偏差の中で 68 % の確率で収まり、 ± 2 標準偏差の中で 95 % の確率で収まり、 ± 3 標準偏差の中で 99 % の確率で収まることを前提としている。標準偏差 σ は、 N は期間、 p はレート、 \bar{p} は期間中の平均のレートとすると、

$$\sigma = \sqrt{\frac{\sum_{i=1}^N (p_i - \bar{p})^2}{N-1}} \quad (2.4)$$

この指標を使った取引ルールの例は、レートが -2σ を超えたときに統計上ではこのあと平均付近に戻る可能性が高いことから買いサインとなる。同じように、レートが 2σ を超えたときに統計上ではこのあと平均付近に戻る可能性が高いことから売りサインとなる。

MACD

この指標では、「MACD」、「シグナル」の 2 つの指標を用いて取引ルールを作るが、「MACD」は「ShortEMA」、「LongEMA」の 2 つの指標から算出される。MACD は以下のように求めることができる。

$$MACD = ShortEMA - LongEMA \quad (2.5)$$

また、シグナルは MACD の指数平滑移動平均により求めることができる。

$$signal = MACD_{t-1} + \alpha(p_t - MACD_{t-1}) \quad (2.6)$$

ここで、 α は式 2.3 と同じである。取引ルールとしては、MACD がシグナルを上から下へ抜ける時が売りサインとなり、MACD がシグナルを下から上へ抜けている時が買いサインとなる。また、MACD の値がプラスの値からマイナスの値に変換したときに下降トレンドとなり売りサインとなり、MACD の値がマイナスの値からプラスの値になったときは上昇トレンドとなり買いサインとなる。

ストキャスティクス

現在の市場に対して一定期間の変動幅に基づいて、売られすぎているか買われすぎているかを判断するための指標である。この指標では、「%K」、「%D」、「%SD」の3つの指標を算出する。

m 日間の %K は、 n 日間の %K の単純移動平均、%SD は n 日間の %D の単純移動平均とすると、

$$\%K = \frac{p_t - p_{min}^m}{p_{max}^m - p_{min}^m} \times 100 \quad (2.7)$$

また、%D は、 n 日間の %K の単純移動平均、%SD は n 日間の %D の単純移動平均となる。

%K と %D を用いた取引ルールの例は、%K・%D が共に 20% 以下の時に、%K が %D を下から上抜いた時が買いサイン、%K・%D が共に 80% 以上の時に、%K が %D を上から下抜いた時が売りサインとなる。

同様に、%D と %SD を用いた取引ルールの例は、%D・%SD が共に 20% 以下の時に、%D がスロー %D を下から上抜いた時が買いサイン、%D・%SD が共に 80% 以上の時に、%D がスロー %D を上から下抜いた時が売りサインとなる。

§ 2.3 ストラテジーテスターにおけるバックテストと最適化

バックテストとは、自分が使っている売買ルールが有効であることを確認するために、ツールを使って過去の相場でシミュレートすることである [4]。主にシステムトレードの分野で使われることの多い言葉だが、裁量トレードの分野でも、売買ルールの有効性を確認するためにバックテストが行われることがある。

バックテストを行うことで感情や長時間の市場の監視や精神的な疲れといった精度を下げる要因を一切排除して、バックテストに利用されている売買ルールの有効性を確認することができる。

売買ルールに用いているインジケータの中には、計算する際に使用する期間等のパラメータを設定する必要がある。通常であればあらかじめ設定したパラメータでインジケータの計算を行い売買ルールに適用するが、実際にはそのパラメータより最適なパラメータが存在するかもしれない。そこで本実験ではバックテストを行うことで最適なパラメータのチューニングを行い、選択されたパラメータを使用して自動売買に適用する。

Python によるバックテスト

Backtesting.py は Python で為替のバックテストを行うことができるライブラリである。ヒストリカルデータを設定して、バックテストを実行し、分析結果を見るというバックテストの基本的な部分がシンプルに作成されているうえに、結果が簡単に一覧表示できるため初心者でも扱いやすいライブラリになっている。

ヒストリカルデータが格納されたデータフレーム、売買ルールが記述されたストラテジークラスを用意するだけで、ヒストリカルデータの期間におけるバックテストを行うことができる。また、ストラテジークラス内に存在する変数の範囲を指定することで、その範囲内で 1 番利益が出る変数の値を最適化して求めることができる。

インジケータの選択

§ 3.1 曜日時間帯による変動の特徴の抽出

本研究では、市場の時間帯における変動の特徴を抽出するとともに、得られた特徴からどのような時間帯に取引を行うとより利益が増えやすいのかということを適用する。

今回の実験では2021年11月1日から2021年12月1日の円ドル為替市場のデータを用いることで実際に曜日や時間帯にどれくらいの違いが出るのかを検証した。グラフにおいて表されているのは横軸が時間であり、縦軸がその時間に行われたトレード回数である。まずどの曜日にも共通しているのがある時刻の付近の時刻はトレード回数が他の時刻と比較して増えている点である。FXは24時間取引することが可能と自由度が高いが、これは東京、香港、シンガポール、ロンドン、ニューヨークといった為替市場開かれており平日であれば常にどこかの市場が動いているからである。グラフにおいてトレード回数が増加している時刻はこれらの市場が開くタイミングであり、最も活発になるニューヨークとロンドン市場が動いている時刻は全ての曜日において取引が活発になっていることがグラフから分かる。

なぜニューヨーク市場やロンドン市場が開くと取引回数が増えるのかというと、市場自体の規模が大きいためです。そのため取引に参加するトレーダの数が他の市場に比べて多いことにより取引回数も増加することが分かっています。また曜日別の特徴で行くと、月曜日の早い時間帯では取引自体が行われていないことが分かる。これはFXは土曜日と日曜日は全ての市場が閉じているからである。また週末に起きた「国際会議などの経済イベント」「中央銀行総裁や政治家などの主要人物の発言」「経済状況」「事件・事故」の影響結果が集約されて為替に反映されるため、大きなニュースがあった場合、早朝6時の時間帯に、機関投資家の取引が始まるため、市場の大きな変動が起こるとともに取引回数も増加する。

このように各市場の開く時間は市場の変動が大きくなることから利益の拡大が狙えることから、それらの時間帯に対して普段よりもトレードの回数を増やす自動売買を考える。

§ 3.2 市場データの選択と活用

本研究では実際に取引において使用する円ドル為替市場と日経平均株価、ニューヨーク・ダウ＝ジョーンズ平均株価、金、原油の市場との関係をグレンジャー因果性分析によって調査する。しかしグレンジャー因果性検定は様々な制約があるためこの節ではグレンジャー因果性検定を行うにあたって必要となるデータの処理について記述する。

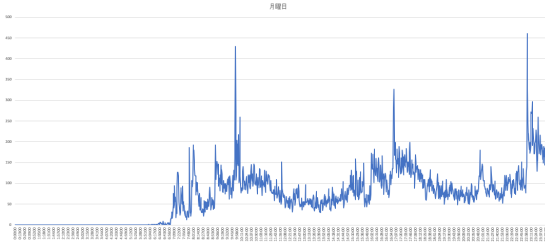


図 3.1: 月曜日

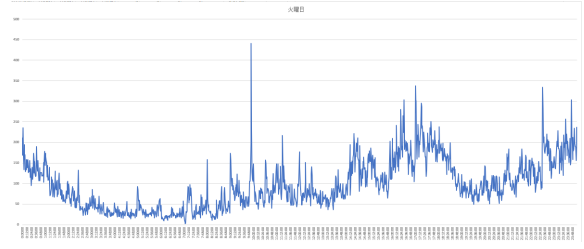


図 3.2: 火曜日

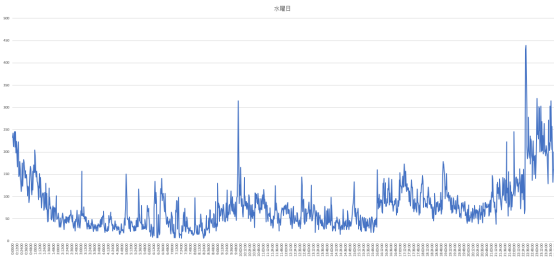


図 3.3: 水曜日

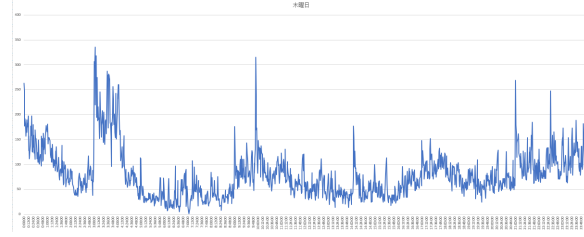


図 3.4: 木曜日

グレンジャー因果性検定を行うにあたって必要となるのが VAR モデルである。冒頭でも記述しているが VAR モデルとは自己回帰モデルを多変量に拡張したものである。VAR モデルが使用される主な目的としては主に 2 つで、1 つは複数の変数を用いて予測精度の向上を図ることであり、もう 1 つは変数間の動学的関係の分析を行うことである。特に 2 つ目の動学的関係の分析においてはデータに対して VAR モデルを適用することで、グレンジャー因果性分析やインパルス応答関数、分散分析といったような分析が使用可能になる [3]。

VAR (ρ) モデルを \mathbf{y}_t を定数と自身の ρ 期の過去の値に回帰したモデルとすると

$$\mathbf{y}_t = \mathbf{c} + \Phi_1 \mathbf{y}_{t-1} + \cdots + \Phi_\rho \mathbf{y}_{t-\rho} + \varepsilon_t, \varepsilon_t \sim W.N.(\Sigma)$$

というモデルである。ここで \mathbf{c} は $n \times 1$ 定数ベクトルであり Φ_i は $n \times n$ 係数行列である。具体的に x_t と y_t という 2 変数で VAR モデルを作成した場合以下の式のように表される。

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + a_3 y_{t-1} + a_4 y_{t-2} + u_{xt}$$

$$y_t = b_1 x_{t-1} + b_2 x_{t-2} + b_3 y_{t-1} + b_4 y_{t-2} + u_{yt}$$

この時のラグ次数は 2 であり、 u は攪乱項である。

しかしグレンジャー因果性検定を行うにあたって、単位根を持つ時系列データ間で回帰モデルや同時方程式モデルなどを作ると、互いに無相関であっても有意な t 値や決定係数が観測されるような見せかけの回帰が起こる可能性があるためこのような非定常な変数を用いた推計は経済時系列分析では意味のないものになる。そこで差分系列に直し共和分検定で共和分で円ドル為替市場と共和分の関係にないことを検定により決定された市場データに対してグレンジャー因果性検定を行っていく。

§ 3.3 因果関係分析による他市場の活用

使用するデータに対して処理を行った上でグレンジャー因果性検定を行う。グレンジャー因果性検定は時系列データだけから因果性の有無を判断できる概念があれば便利である，そのような考えをもとに1969年に，何の理論にも基づかない予測を基準とする因果性が提案された。グレンジャー因果性検定は以下のように定義されている [5]。

定義 4.1 (グレンジャー因果性検定)

現在と過去の x の値だけに基づいた将来の x の予測と，現在と過去の x と y に基づいた将来の x の値を予測して比較して，後者の平均二乗誤差 (Mean Squared Error: MSE) の方が小さくなる場合， y_t から x_t へのグレンジャー因果性が存在するといわれる。

グレンジャー因果性検定は，時系列モデルにおいてある変数 x が他の変数 y に影響を及ぼす，より具体的には他の条件を一定として y の過去の値が x の変動について説明力を持つということであり，論理的な意味での因果性とは意味が異なる [6]。またグレンジャー因果性検定は複数の変数が存在する一般的な場合にも簡単に拡張できる。

定義 4.2 (一般的なグレンジャー因果性検定)

x_t と y_t をベクトル過程とする。また Ω_t とし Ω_t から現在と過去の y を取り除いたものを $\tilde{\Omega}_t$ とする。この時， $\tilde{\Omega}_t$ に基づいた将来の x の予測と， Ω_t に基づいた将来の x を比較して，後者の MSE の方が小さくなる場合， y_t から x_t へのグレンジャー因果性が存在するといわれている。ここで，MSE の代償は行列の意味での大小であることに注意されたい。

これらの定義に基づいて n 変量 VAR(p) におけるグレンジャー因果性検定を行うと以下のような流れになる。

n 変量 VAR (p) におけるグレンジャー因果性検定の手順

1. VAR モデルにおける y_{kt} のモデルを最小二乗回帰 (Ordinary Least Squares regression: OLS) で推定し，その残差平方和を SSR_1 とする。
2. VAR モデルにおける y_{kt} のモデルに制約を課したモデルを OLS で推定し，その残差平方和を SSR_0 とする。
3. F 統計量を

$$F \equiv \frac{SSR_0 - SSR_1/r}{SSR_1/(T - np - 1)}$$

で計算する。ここで， r はグレンジャー因果性検定に必要な制約の数である。

4. rF を $\chi^2(r)$ の 95% 点と比較し， rF の方が大きければ，ある変数から y_{kt} へのグレンジャー因果性は存在し，小さければグレンジャー因果性は存在しないと結論する。

このような検定を行い，日経平均株価，ニューヨーク・ダウ＝ジョーンズ平均株価，金，原油の市場が円ドル為替市場を予測するにあたって使用できることが確認出来たら，円ドル為替

Granger causality F-test. H_0: us30 does not Granger-cause tick. Conclusion: fail to reject H_0 at 5% significance level.			
=====			
Test statistic	Critical value	p-value	df

0.5199	1.298	1.000	(68, 34636)

Granger causality F-test. H_0: jp225 does not Granger-cause tick. Conclusion: fail to reject H_0 at 5% significance level.			
=====			
Test statistic	Critical value	p-value	df

0.8902	1.298	0.728	(68, 34636)

Granger causality F-test. H_0: gold does not Granger-cause tick. Conclusion: fail to reject H_0 at 5% significance level.			
=====			
Test statistic	Critical value	p-value	df

0.7074	1.298	0.968	(68, 34636)

図 3.5: グレンジャー因果性検定

市場とその他の市場との相関を求めることで得られた相関係数を利用する。実際に 2021 年 10 月 25 日から 2021 年 10 月 26 日の円ドル為替市場に対して日経平均株価，ニューヨーク・ダウ＝ジョーンズ平均株価，金の 3 つの市場が因果性を持つのかを確認する（図 3.5 参照）。

提案手法

§ 4.1 機械学習のためのデータセット作成

従来研究の多くは、自身が作成したシステムを過去のヒストリカルデータに適用して予測精度の向上について検証しているものが多い。しかし、過去のデータはあくまでも過去のデータであり、これから先の未来の市場で同じ値動きが発生する可能性は低いと言える。

本研究では、リアルタイムで動いている市場から最新の Tick データを取得・蓄積した後指定時間足でリサンプリングを行うことで、より直近の値動きから作成されたヒストリカルデータを使用する。その後、最適化したヒストリカルデータを用いてインジケータのパラメータを最適化する。ヒストリカルデータの作成からインジケータのパラメータ最適化は以下のような流れで行う。

1. Tick データ更新ごとに MT5 から Tick データの取得
2. 1 秒間に一回だけ Tick データを保存、データフレームに格納
3. 指定秒足で OHLCV の形にリサンプリングし、CSV ファイルに保存
4. 作成したヒストリカルデータを用いてインジケータのパラメータ最適化を行う
5. 最適化したパラメータを CSV ファイルに書き出し保存

まず最初に MT5 から Tick データを取得してくる。取得は Python を使用して行う。MT5 には Python を使ってデータをやり取りするためのモジュールが用意されているため今回はそれを利用した [7]。while 文で永続的にプログラムは動き続け、該当部分のコードに到達するたびに、指定した通貨ペアの Tick の値を MT5 から取得する。

また、Tick データは同じ時間に複数のデータが送られてくる場合がある。このような事態を避けるために、Tick データは永続的に取得はするが、保存は一つ前のデータから秒数の部分が変更したタイミングだけ行うように設定している。また、取得した Tick データ内に含まれる時間のデータは UNIX 時間であるため、保存する前に datetime 型に変換してからデータフレームに保存している。

さらに、今回の手法では取得してきた Tick データに含まれている BID と ASK の値の平均を取り、その値を今後使う価格の値としている。両方の値を使用しようとした場合、ヒストリカルデータの作成やパラメータ最適化、この後説明するルール選択の処理がそれぞれ 2 倍になってしまうため、処理時間も増えてしまう。本研究ではルール選択にかかる時間はなるべく少なくしたいため、BID と ASK の値の平均値を使うことで 1 つのヒストリカルデータとしている。

また、インジケータの計算には価格のデータのほかに Volume の値も必要なものも存在する。そのため、保存する Tick データに含まれる Volume を一緒に保存する。つまり、保存先のデータフレームには、Time, Price, Volume の 3 つが保存されていく。

このままでは保存したデータフレームはプログラムが動いている間は溜まり続けるが、プログラムが再度動きなおした際には集めた Tick データは全てリセットされてしまう。それを避けるために Tick データが増えるたびに CSV ファイルに保存し、プログラムの 1 番最初でその CSV ファイルを読み取りに行くことで、今まで貯めたデータを継続して使用できるようにしている。

保存したデータフレームには価格の値は 1 つだけである。しかし、インジケータの計算に使用するデータセットは OHLC 型もしくは OHLCV 型である必要がある。そこで、データフレームの中身を指定時間ごとにリサンプリングを行い、OHLCV 型に変換したデータフレームを作成したのちに CSV ファイルに保存する。

ここで、リサンプリングに指定する時間は 10 秒, 30 秒, 1 分, 3 分, 5 分, 10 分, 15 分の 7 種類用意し、それぞれ別の CSV ファイルに保存する。これは、使用時に様々な時間足のデータセットの中から予測に使うものを選ぶことができるようにするためである。

また、保存したデータセットはリサンプリングが行われるたびに最新のものに更新するために書き込みが行われる。また、この後のインジケータの計算に使用するために頻繁に読み込みが行われる。今回複数のインジケータを同一のデータセットを使って計算するため、データセットを保存したファイルが 1 つだけだと読み書きの際に衝突が起こってしまう恐れがある。これらを避けるために、今回複数の CSV ファイルをインジケータの数だけ作成し、全て同じデータセットを保存しそれぞれ別の CSV ファイルから読み込みを行うことで衝突が起こることを避けている。

取得した Tick データや作成したヒストリカルデータはプログラムが動き続けている間は常に溜まり続ける上、プログラムを再起動してもリセットされることなく続きから蓄積されていく。このままではデータの数が増えすぎて読み書きの処理が重くなってしまう恐れがあるため、あらかじめ指定したデータ数を超えたら古いものから削除するようにしている。

ヒストリカルデータの作成後、そのデータを用いてインジケータのパラメータの最適化を行う。今回は 2.3 章で説明した Python でバックテストが行える Backtesting.py というライブラリを使用する。また、インジケータの計算には TA-Lib というライブラリを使用する。TA-Lib は、複数の言語で使用可能なテクニカル指標の分析ツールで、ヒストリカルデータとそれぞれのインジケータに必要な期間等の数値を与えるだけでインジケータの計算をすることができる。

各インジケータのバックテストの際にはそれぞれのインジケータだけでなく、同時に Average True Range (ATR) というインジケータも使用する。ATR は価格変動の度合いを計測するためのインジケータで、ATR の値が高ければ取れる利幅も大きくなる一方で大きな損失を被るリスクも増大し、ATR の値が低ければ損失を被るリスクは低減するが取れる利幅も少なくなるといった判断ができる。

バックテストの際にはそれぞれのインジケータごとに売買ルールを設定し、売買タイミングが発生したらオーダーを送る。その際には利確と損切りの幅の値も同時に送り、それぞれの幅分の価格が動いた時点で決済を行っている。今回のシステムではこの利確と損切りの幅に ATR を利用する。

また、ATR の値をそのまま利確と損切りの幅に使用してしまうと幅が小さすぎて決済するまでの時間が早くなり、利確したとしてもスプレッドの方が大きくなってしまう。それを避けるために、ATR の値を数倍した値を利確と損切りの幅に指定し、その倍率も最適化を行う。つまり、今回のシステムで最適化を行うパラメータは以下の 3 点である。

1. 各インジケータの計算に必要な期間（複数必要な場合もある）
2. ATR の計算に必要な期間
3. ATR の倍率

最適化が終了したらその結果得られた最適なパラメータを CSV ファイルに保存する。また最適なパラメータを使用した時の各インジケータの数値、また Tick データにおいてある一定の期間で市場の変動が一定以上あったのかを判断したデータを OHLCV データの時間足と共に csv ファイルに保存する。インジケータの最適化を行うプログラムは各インジケータごとに用意し、それぞれのインジケータごとに最適化が終わり次第新しいパラメータを更新していく。

§ 4.2 機械学習によるルール作成

機械学習に使用するデータは 4. 1 章の工程で作成した、最適なパラメータを使用した時の各インジケータとそれらの数値に対応する OHLCV データの時間足が保存されたものである。これらのデータに対して多層パーセプトロン（Multilayer perceptron: MLP）を使用する。

MLP は複数の形式のニューロンが多層に接続されたネットワークを指す。また MLP は学習のために誤差逆伝播法と呼ばれる教師あり学習手法を利用する。

今回の実験において用いる教師データは、過去のバックテストにおいて最適なパラメータを使用した時の各インジケータの数値、また Tick データにおいてある一定の期間で市場の変動が一定以上あったのかを判断したデータと OHLCV データの時間足が一緒になった csv ファイルです。これらを入力することでこれから先の一定期間で市場の変動が一定以上あるのかを出力し、それらを考慮して自動売買を行う。

§ 4.3 提案手法のアルゴリズム

最後に、本研究で提案した複数ルールの多目的効用最大化を考慮した自動売買システムのアルゴリズムについてまとめる（図 4.1 参照）。

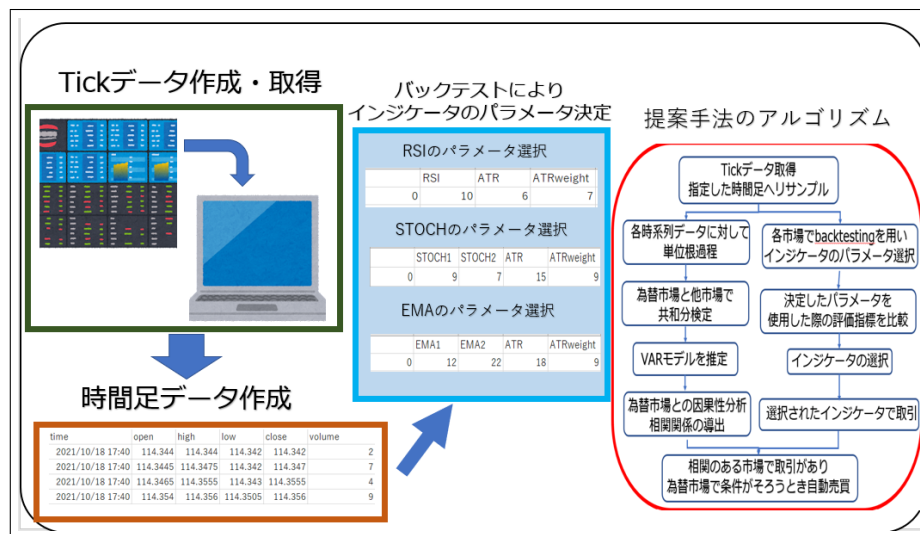


図 4.1: 提案手法の流れ

[Step1] データ取得とヒストリカルデータ作成

Python を用いて MT5 から Tick データを取得してくる。Tick データは永続的に取得はするが、保存は一つ前のデータから秒数の部分が変わったタイミングだけ行う。保存したデータフレームには価格の値は1つだけであるが、インジケータの計算に使用するデータセットは OHLC 型もしくは OHLCV 型である必要があるため、データフレームの中身を指定時間ごとにリサンプリングすることで、OHLCV 型に変換したデータフレームを作成したのちに CSV ファイルに保存する。10 秒、30 秒、1 分、3 分、5 分、10 分、15 分の 7 種類のうち指定した時間のリサンプリングを行い、csv ファイルに保存する。

[Step2] インジケータのパラメータ最適化

作成したヒストリカルデータを用いてインジケータのパラメータの最適化を行う。最適化には Backtesting.py というライブラリを使用し、各インジケータの計算に必要な期間、利確と損切りの幅に利用する ATR の計算に必要な期間、ATR の倍率のパラメータを最適化する。最適化が終了したらその結果得られた最適なパラメータを CSV ファイルに保存する。インジケータの最適化を行うプログラムは各インジケータごとに用意し、それぞれのインジケータごとに最適化が終わり次第新しいパラメータを更新していく。これらの流れを円ドル為替市場だけでなく、自動売買に用いる全ての市場で行う。

[Step3] 円ドル為替市場に対するグレンジャー因果性検定

各市場の指定した時間足のリサンプリングデータから Close のデータを取り出す。それらのデータに対して 3 章で記したように、単位根検定、共積分過程を行いグレンジャー因果性検定に使用できる市場のデータのみで検定を行う。それらから実際にグレンジャー因果性因果が存在することが確認出来たら、相関係数を求める。グレンジャー因果性の有無、相関係数を csv ファイルに保存する。

[Step4] 最適なパラメータ, グレンジャー因果性を用いての売買選択

step2で抽出された最適なパラメータを用いて現在の市場に売買シグナルが出ているかを確認する。これらの処理は円ドル為替市場以外の市場においても行う。売りのシグナルが出ている場合は-1, 買いのシグナルが出ている場合は1, 売買シグナルが出ていない場合は0をcsvファイルに保存する。そのcsvファイルを相関に基づいて分けていく。円ドル為替市場の場合はシグナル通りの売買をしてよいが, 他市場において相関がマイナスの場合, シグナルとは逆の売買を行わなくてはならない。そのようにして売買を分けたうえで円ドル為替市場に出ているシグナルとその他の市場が示すシグナルが同じでありかつ数が一定の値を上回った場合その円ドル為替市場に出ているルールを選択する。

[Step5] 最適ルールでの自動売買

最適なルールが得られたらそのルールに則って自動売買を行う。取引はMT5のデモ口座を使用して行う。ヒストリカルデータと最適なパラメータから計算したインジケータの値を利用して, 売買タイミングになったらPythonを用いて売買オーダーを送る。売買オーダーを送る際には, バックテストを行った際と同じ条件になるように設定し, エントリー後に利確と損切りの幅に達したら決済を行う。

数値実験並びに考察

§ 5.1 数値実験の概要

数値本研究では、4.3章で説明したように、データ取得とヒストリカルデータ作成、インジケーターのパラメータ最適化、円ドル為替市場に対するグレンジャー因果性検定、最適なパラメータ、グレンジャー因果性を用いてのルール選択、最適ルールでの自動売買の5つの工程がある。

まず、Pythonを用いてMT5からTickデータ収集する。Tickデータは取得したTickデータの秒の部分が変わっていたらデータフレーム保存していく。今回は収集したTickデータの中から時間、価格、ボリュームを抜き出して保存している。

価格はBIDとASKの2種類があるが、Backtesting.pyではそれぞれにヒストリカルデータを作り、バックテストをする必要があるが、本研究では売買の選択をするにあたって実際の市場とバックテストに使用するデータとの違いを最小にするためにそれぞれの平均値を使うことで1つのヒストリカルデータとしている。

その後、保存したTickデータを指定した時間足でリサンプリングし、CSVファイルに保存する。

ヒストリカルデータを作成後、各インジケーターのパラメータ最適化を行う。インジケーターの最適化はそれぞれインジケータごとに別々のプログラムで実行し、それぞれ最適化が完了するごとに最適なパラメータを更新していく。最適化にはBacktesting.pyを、それぞれのインジケーターの計算にはTA-Libを使用している。トレンドラインに関してはTA-Libでの計算ができないのでプログラムを作成して計算を行った。今回使用したインジケーターは以下の7つである。

1. EMA
2. ボリンジャーバンド
3. MACD
4. 相対力指数 (Relative Strength Index: RSI)
5. ストキャスティクス
6. 方向性指数 (Directional Movement Index: DMI)
7. テンドライン

全てのインジケーターの売買ルールで、オーダー時の利確と損切り幅にはATRというインジケーターを数倍したものをを用いている。利確の幅はバックテストによって求められた

幅を追加している。最適化する際の幅は、一般的に使われている期間から飛び幅2ずつで、前後3つずつ入るように設定している。ただし、MACDとULTOSCは変数の数が多いより多く、最適化に時間がかかるため、前後2つずつ入るような範囲の設定している。

今回は初期所持金100,000円、手数料0.2pipsに設定してバックテストを行い、最適なパラメータを決定する。また、評価指標のうちEquity Finalが最大になるように最適化を行う。最適化が完了したら、得られたパラメータをインジケータごとに別々のCSVファイルに保存する。

また作成されたOHLCVデータを使用して、円ドル為替市場と日経平均株価、ニューヨーク・ダウ＝ジョーンズ平均株価、金、原油の市場とのグレンジャー因果性を調査する。各市場のOHLCVデータのCloseの値を使用する。最初に単位根検定を行う、その後共和分検定を行い、円ドル為替市場と共和分の関係にあると分かったものはグレンジャー因果性検定に使用しない。

グレンジャー因果性検定を行った後に、円ドル為替市場との相関係数を求める。グレンジャー因果性の有無と相関係数をcsvファイルに保存する。

得られた最適なパラメータをリアルタイムで動いている市場に適用して市場に対して売買のシグナルが出ているかを確認する。この時円ドル為替市場のみならず、日経平均株価や原油などの他の市場に対しても同様のことを行う。売りのシグナルが出ている場合は-1、買いのシグナルが出ている場合は1、売買シグナルが出ていない場合は0をcsvファイルに保存する。

売買シグナルがまとめられたcsvファイルとグレンジャー因果性と相関係数が保存されているcsvファイルを使用することで、売買シグナルを相関と因果性の有無について分けていく。円ドル為替市場はシグナル通りの売り買いと判断し、他市場においては因果性が有り、相関係数が-0.6以下の時シグナルとは逆の売買と判断する。因果性が有り、相関係数が0.6以上の時シグナル通りの売り買いと判断する。

そこで円ドル為替市場に出ている売買シグナルと同様のシグナルが因果性を持つ他市場でも4個以上出ているかつ逆の判断をしているシグナルより多い場合取引を行う。

1. エントリーする銘柄
2. ロット数
3. エントリーする価格
4. 利確幅
5. 損切り幅
6. スリップページ

今回は銘柄はUSDJPY、ロット数は0.1、スリップページは20で設定している。価格はそのときの最新のTickデータを取得し、売りでエントリーする際にはBIDの価格で、買いでエントリーする際にはASKの値を使用する。利確と損切りの幅はそれぞれのインジケータで最適化されたATRの値と倍率をかけたものを使用する。ただし、利確幅は一般に損切り幅よりもスプレッドの値分大きくする。今回はバックテストの際にスプレッドを0.2pipsに設定したので、自動売買の際も損切り幅から0.2pips分大きくした値を使用する。

実際にこのシステムを動かした結果を以下に示す。Tickデータは2021年11月1日から取得し始め、実際にその後のシステムを動かす際に必要なデータ数が収集できるようにした。

表 5.1: 実験結果

時刻	オーダー	収支 (円)
12/20 04:36:06	sell	+590
12/20 04:36:06	sell	+590
12/20 06:31:01	sell	+590
12/20 13:42:46	buy	+590
12/20 16:20:19	sell	+540
12/20 16:20:19	sell	+540
12/21 12:03:55	sell	-300
	総収支	+3150

また、インジケータの最適化プログラムはインジケータの計算に必要な期間以上のデータ数が蓄積した段階から動かし始め、その後は随時更新し続けて常に最適なパラメータを更新し続ける。またグレンジャー因果性検定の結果も更新をし続ける。

上記のプログラムを動かした後、自動売買のプログラムを起動することで自動売買が始まる。今回は自動売買のプログラムを 12/20 4:30 から 12/21 12:30 まで動かした。その際の結果を表 5.1 に示す。実験期間中の取引回数は 7 回、勝率は約 86%で収支は+3150 円であった。

§ 5.2 実験結果と考察

今回の実験結果より、提案手法で利益を上げることが出来た。勝率が 9 割近くあることから、市場から得られるインジケータの値と実際に取引をしない市場からの影響を考慮した今回の手法は最適な売買を行うことができると考える。

また一回の取引における収支を見ると、バックテストで最適な利確幅を追加することが出来ているので、勝率が 5 割であったとしても収支的にはマイナスにならずにプラスになることが分かる。

今回の実験は約 1 日半の間プログラムを動かしたが、取引回数自体が 7 回ととても少なかった。これは市場間分析においてグレンジャー因果性を有する市場が存在しないと円ドル為替市場でシグナルが出ても売り買いはしないことから取引をするにあたって条件が大変厳しいものであったと考えられる。しかしそれらの条件があったからこそ今回のような結果が得られるとも考えられる。

このように取引回数が少ないのを改善するには、今回の実験で指定した条件を現在のものより取引回数が増えるように変更するか、他の売り買いの判断も行えるような判断方法と併用して使用することで取引回数の改善ができるのではないかと考える。

おわりに

従来研究で行われていた，ある市場が他の市場に与える影響を予測するという研究を自動売買に組み込めた．

謝辞

本研究を遂行するにあたり，多大なご指導と終始懇切丁寧なご鞭撻を賜った富山県立大学工学部電子・情報工学科情報基盤工学講座の奥原浩之教授，António Oliveira Nzinga René 講師に深甚な謝意を表します．最後になりましたが，多大な協力をしていただいた研究室の同輩諸氏に感謝致します．

2022 年 2 月

木下 大輔

参考文献

- [1] 高橋 和秀, “先物外国為替市場の予測性についての一考察”, 関西学院商学研究, Vol. 62, pp. 95-116, Mar. 2010.
- [2] 温井 慧, 高木 徹 “AI を用いた FX システムトレードの提案”, 経営情報学会 全国研究発表大会要旨集, pp. 1-30, Oct. 2019.
- [3] 内藤 友紀, “2008～09 年の日本における株式価格の下落について : VAR モデルによる要因分析”, 関西大学経済論集, Vol. 60, No. 1, pp. 1-18, Jun. 2010.
- [4] “【FX 用語】バックテストとは? 言葉の意味とその重要性について”, <https://fx-square.com/backtest/>. 閲覧日 2021.12.25.
- [5] 沖本 竜義, “経済・ファイナンスデータの計量時系列”, 朝倉書店, 2010.
- [6] 宮崎 慧 星野 崇, “階層ベイズ動的モデルによるブランドスイッチングの分析: グレンジャー因果性検定の利用”, マーケティング・サイエンス, Vol. 21, No. 1, pp. 11-35, 2013.
- [7] “MT5 (メタトレーダー 5) のバックテストの方法 — OANDA FX/CFD Lab education (オアンダ ラボ) ”, https://www.oanda.jp/lab-education/blog_mt5/auto_trading/mt5backtest/, 閲覧日 2021.12.25.