

高頻度データに対する 多目的効用最大化のための ストラテジーの自動チューニング

大谷 和樹

富山県立大学 情報基盤工学講座

1. はじめに
2. 高頻度データ収集
3. 直交表にもとづくロバスト設計
4. 提案手法
5. 実験結果ならびに考察
6. おわりに

はじめに

高頻度データ収集

直交表にもとづく
ロバスト設計

提案手法

実験結果ならびに
考察

おわりに

背景

経済活動の活発化に伴い、金融市場の規模は拡大し金融市場のメカニズムに関する研究の重要性が増している。しかし、金融市場は複雑な要因が絡み合っているため全容を明らかにすることは困難である。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。

目的

過去の研究の多くは複数ルールは適用していたとしても、効率的にルールを選ぶ方法は組み込まれていない。そこで、本研究ではリアルタイムで取得した Tick データから指標を算出し、より最適な売買ルールを自動的に選択する金融マーケット予測手法の開発を目的とした分析手法を考える。

MetaTrader 5(MT5)

外国為替及び為替市場におけるテクニカル分析及び取引業務を行うトレーダー向けの無料アプリケーションである。

Python を使って MT5 から Tick の情報を取得したり売買の指令を送ることができるため、取得した Tick の情報を利用した自動売買を行うことができる。

はじめに

高頻度データ収集

直交表にもとづく
ロバスト設計

提案手法

実験結果ならびに
考察

おわりに

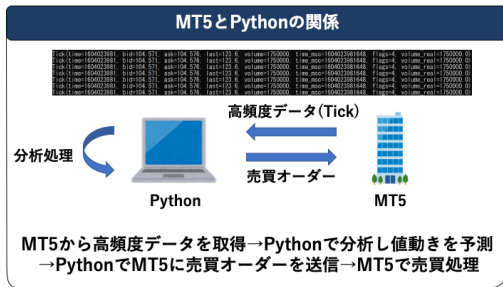


Figure 1: MT5 と Python の関係

テクニカル分析

過去と現在の価格や出来高などのデータからテクニカル指標を算出することによって傾向を把握し、これからの値動きについて未来の価格の予測を行う。

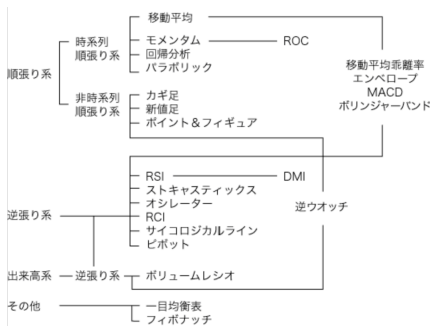


Figure 2: テクニカル分析に使われる指標

はじめに

高頻度データ収集

直交表にもとづく
ロバスト設計

提案手法

実験結果ならびに
考察

おわりに

インジケータ

テクニカル分析に使用するために、為替レートの時系列情報を様々な計算で加工して売買の判定に利用する指標のことを指す。特定期間内の平均レートや相場の方向感を数値化したデータなどが例として挙げられる。

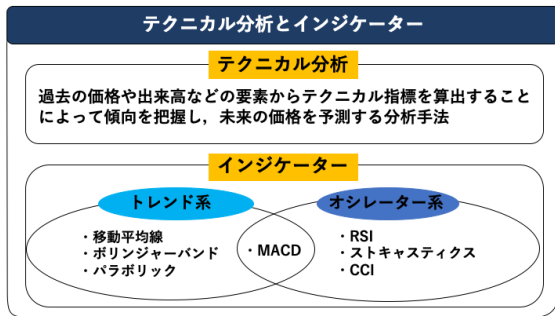


Figure 3: インジケータ

ロバスト設計

様々なノイズ（誤差因子）に強い設計を行う強力な設計最適化手法.

パラメータ設計の基本的な考え方は、ばらつきの原因となる誤差因子をコントロールするのではなく、設計に有効な制御因子と誤差因子の交互作用（誤差因子の影響がなるべく小さくなるような制御因子の水準条件）を見つけることにより誤差因子の影響を減衰させようとするものである.

パラメータ設計は、制御因子の水準変更のみでばらつきの低減を図れるという経済的かつ効果的な方法であるため、ものづくりの設計開発の現場を中心に利用されている.

直交表

すべての列において数値の組み合わせが均衡，列同士の相関が全くない，列同士の成分を掛け合わせた値を持つ列が存在するといった性質を持っている。

直交表の列に要因を割り付け，それぞれの水準を対応させることで調査条件を設定することで計画行列と見ることができる。

直交表 $L_8(2^7)$

No.	1	2	3	4	5	6	7
1	1	1	1	1	1	1	1
2	1	1	1	2	2	2	2
3	1	2	2	1	1	2	2
4	1	2	2	2	2	1	1
5	2	1	2	1	2	1	2
6	2	1	2	2	1	2	1
7	2	2	1	1	2	2	1
8	2	2	1	2	1	1	2

- ・ 2水準の要因を7個まで扱える
- ・ 多元配置の $128(2^7)$ 通りの組み合わせが8通りの組み合わせで確認できる

直交表 $L_9(3^4)$

No.	1	2	3	4
1	1	1	1	1
2	1	2	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

- ・ 3水準の要因を4個まで扱える
- ・ 多元配置の $81(3^4)$ 通りの組み合わせが9通りの組み合わせで確認できる

Figure 4: 直交表

実験計画法

直交表を用いた要因計画は、実験計画法のうちの一つである。
実験計画法では、直交表を回帰分析し、その測定結果 y_{ijk} をモデル化する (i, j, k はパラメータの水準を表す)。
これにより、測定結果における各要因の影響度を求めることができる。

直交表とモデル式

No.	因子1	因子2	因子3	結果
1	1	1	1	1
2	1	1	2	2
3	1	2	1	3
4	1	2	2	4
5	2	1	1	5
6	2	1	2	6
7	2	2	1	7
8	2	2	2	8

パラメータが3つの場合のモデル式

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\beta\gamma)_{jk} + (\alpha\gamma)_{ik} + (\alpha\beta\gamma)_{ijk} + \varepsilon_{ijk}$$

上記のような
直交表を
分散分析

μ : パラメータによらない成分
 $\alpha_i, \beta_j, \gamma_k$: パラメータA,Bそれぞれによる効果
 $(\alpha\beta\gamma)_{ijk}$: パラメータ間の交互作用による効果
 ε_{ijk} : 測定の誤差

Figure 5: モデル式

はじめに

高頻度データ収集

直交表にもとづく
ロバスト設計

提案手法

実験結果ならびに
考察

おわりに

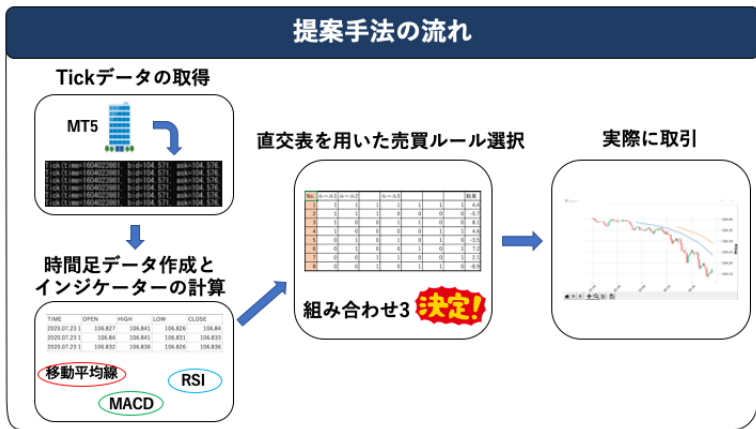


Figure 6: 提案手法の流れ

MT5 経由で Tick データを取得する.

取得タイミングは値動きがあったタイミング→一定じゃない
指定した間隔 (1s, 10s, 1m 等) で resample を行い, 同時にインジケーターの計算に使えるように ohlc の形で保存.

	time	price
0	2020/12/11 4:51	104.283
1	2020/12/11 4:51	104.283
2	2020/12/11 4:51	104.283
3	2020/12/11 4:51	104.283
4	2020/12/11 4:51	104.283
5	2020/12/11 4:51	104.283
6	2020/12/11 4:51	104.283
7	2020/12/11 4:51	104.283
8	2020/12/11 4:51	104.283
9	2020/12/11 4:51	104.283
10	2020/12/11 4:51	104.283
11	2020/12/11 4:51	104.283
12	2020/12/11 4:51	104.283
13	2020/12/11 4:51	104.283
14	2020/12/11 4:51	104.283
15	2020/12/11 4:51	104.283
16	2020/12/11 4:51	104.283
17	2020/12/11 4:51	104.283
18	2020/12/11 4:51	104.283
19	2020/12/11 4:51	104.283
20	2020/12/11 4:51	104.283

Figure 7: Tick データ取得の例

time	open	high	low	close
2020/12/11 4:57	104.2975	104.3005	104.2955	104.3005
2020/12/11 4:57	104.301	104.3015	104.3005	104.3005
2020/12/11 4:57	104.3015	104.3015	104.3005	104.3005
2020/12/11 4:58	104.3005	104.3055	104.3005	104.3055
2020/12/11 4:58	104.305	104.3065	104.3045	104.3055
2020/12/11 4:58	104.3065	104.3065	104.2995	104.3
2020/12/11 4:58	104.3025	104.3025	104.2975	104.2985
2020/12/11 4:58	104.3015	104.3015	104.2995	104.2995
2020/12/11 4:58	104.3	104.301	104.3	104.3005
2020/12/11 4:59	104.301	104.301	104.294	104.2975
2020/12/11 4:59	104.297	104.2975	104.2935	104.2945
2020/12/11 4:59	104.2955	104.2965	104.2945	104.2955
2020/12/11 4:59	104.2945	104.297	104.2945	104.297
2020/12/11 4:59	104.2965	104.2985	104.2965	104.2985
2020/12/11 4:59	104.299	104.3045	104.297	104.304
2020/12/11 5:00	104.3035	104.3325	104.3025	104.325
2020/12/11 5:00	104.3225	104.3265	104.3175	104.3215
2020/12/11 5:00	104.321	104.3345	104.318	104.3295
2020/12/11 5:00	104.329	104.334	104.3245	104.3325
2020/12/11 5:00	104.33	104.3325	104.3225	104.325

Figure 8: resample 後

backtesting.py

オープンソースとして公開されているバックテストのライブラリ。
特徴：処理が高速・Ta-Lib をサポートしている

```
#EMA最適化
print("-----EMA-
bt = Backtest(df_, EMAt, cas
m=list(range(5, 50, 5))
n=list(range(10, 100, 5))
stats = bt.optimize(m=m,n=n,
print(stats)
```

Figure 9: 最適化を行う部分のコード

```
Start          2020-09-24 00:00:00
End            2020-10-01 00:00:00
Duration       7 days 00:00:00
Exposure Time [%] 24.1408
Equity Final [$] 9129.13
Equity Peak [$] 10000
Return [%] -8.70867
Buy & Hold Return [%] 0.118603
Max. Drawdown [%] -8.74479
Avg. Drawdown [%] -8.74479
Max. Drawdown Duration 6 days 21:17:00
Avg. Drawdown Duration 6 days 21:17:00
# Trades 75
Win Rate [%] 0
Best Trade [%] -0.06029
Worst Trade [%] -0.159959
Avg. Trade [%] -0.122062
Max. Trade Duration 0 days 01:40:00
Avg. Trade Duration 0 days 00:23:00
Profit Factor 0
Expectancy [%] NaN
SQN -40.3811
Sharpe Ratio -4.68047
Sortino Ratio -4.68047
Calmar Ratio -0.0139582
_strategy EMA(m=45, n=95)
equity_curve
trades Size EntryB...
```

Figure 10: 最適化した結果

はじめに

高頻度データ収集

直交表にもとづく
ロバスト設計

提案手法

実験結果ならびに
考察

おわりに

直交表を用いたルール選択

鋭意制作中。想定している流れは、

- 1 直交表にもとずいて使うルールを選び、前段階で最適化したパラメータでバックテスト
- 2 最終的な利益の部分だけを抜き出し変数に保存
- 3 結果を比較し最大のものをルールとして選択

実験結果ができたなら載せます.

はじめに

高頻度データ収集

直交表にもとづく
ロバスト設計

提案手法

実験結果ならびに
考察

おわりに

今回の進捗

- ① インジケータ (売買ルール) の数を増加
- ② 複数のインジケータを連続して最適化

今後の課題

- 1 直交表を用いたルール選択のプログラムの作成を急ぐ
- 2 できたプログラムで実際に売買を行う
- 3 インジケータの数 (ルールの数) を増やす
- 4 並行して本論を書き進める