暗号通貨に対する統計的手法による投資戦略の検討

5CS23 小峯谷勇二

2019 年 2 月 4 日

目次

1　研究背景

近年ビットコイン (以降 BTC) をはじめとした暗号通貨が開発されており，それらは各種メディアの露出を機に知名度が向上したと言える．また，BTC 決済に対応した店舗や通信販売も出現しており，BTC を従来の通貨同様決済に使用することが可能になっている．[1]

2　研究目的

BTC の価格は従来の通貨同様常に変動しており，時点ごとの価格が存在する金融系の時系列データである．同等のものには他にも株式・外国為替があり，それに対する統計的な解析及び投機・投資対象としての投資戦略の検討は今日も積極的に行われている．また，BTC も株式・外国為替同様に投資対象として取引が行われている．

BTC と株式・外国為替の決定的な相違点は，取引所ごとのメンテナンスを除けば毎日 24 時間取引が行われている点である．これのメリットとして，前日が市場休場日だった場合の市場開場時 (寄付き) の急激な価格変動が発生しにくいことが挙げられる．

本研究の目的は，BTC の価格データに対して，統計的手法を用いて解析を行い，投資戦略を検討することである．

3　先行研究等

金融系時系列データである株価に対して，データマイニングを利用してその価格を予測するシステムの開発 [2] などが行われている．また，外国為替に対しては，多変量解析による価格の変動を要因を分析する研究 [3] などが行われている．しかし，暗号通貨の価格に対して行われた研究論文は，2019 年 1 月現在ごく少数にとどまっている．従って，本研究では暗号通貨の価格に対して統計的に解析を行う．

4　手法

本研究は大きく分けて以下の手順で行う．

1. BTC 価格データの取得
2. 価格データに対する統計的性質の調査
3. 価格データ解析・解析結果による投資戦略の検討
4. 検討した投資戦略による実データを使用した売買シミュレーション

BTC 価格の取得は，Cryptowatch[4] という API を利用する．本研究においては取引所 bitFlyer(https://bitflyer.com/ja-jp/) の価格データを取得する．API の使用方法は，参考文献の URL を参照のこと．

なお，本研究において，売買シミュレーションに使用する BTC 価格データは以下の通り．

- 範囲: 2018 年 10 月 31 日 12 時 00 分 00 秒から 2018 年 11 月 30 日 23 時 55 分 00 秒まで
- 価格データの最小単位: 5 分ごとの価格

5　開発環境

本研究は，macOS Sierra10.12.6（16G29）上で行っている．

また，本研究では，BTC の価格データの保存・管理をデータベース (以下 DB) で行なっており，本研究で使用しているプログラミング言語は R・python である．本研究において，プログラミング言語 R は，BTC 価格データの統計的性質の調査に使用している．また，プログラミング言語 python は，API による BTC 価格データの取得及び BTC 価格データの解析に使用している．

以下にデータベース及びプログラミング言語の環境を示す．

5.1　DB について

本研究においては，DB は MySQL5.7.21 を使用している．[5]

5.2　R についての開発環境

本研究における，プログラミング言語 R を使用する場合についての環境を示す．

- Version: 3.4.4[6]
- 外部パッケージ
  - 時系列解析パッケージ tseries
  - 時系列データ操作パッケージ urca
  - DB 操作パッケージ RMySQL

5.3　python についての開発環境

本研究における，BTC 価格データの取得にプログラミング言語 python を使用する場合についての環境を示す．

- Version: 3.6.4
- DB 接続用外部パッケージ mysqlclient[7]

続いて本研究における，BTC 価格データの解析にプログラミング言語 python を使用する場合についての環境を示す．

- Version: anaconda3-5.1.0
- テクニカル解析ライブラリ Ta-Lib[8]

6　BTC 価格の統計的性質の調査方法

6.1　調査に使用するデータについて

時系列解析では，解析対象のデータそのもの (原系列データ) の対数差分をとり，その対数差分を解析する場合が多い．

今回解析対象の時系列データである BTC の価格データでは，t 時点での価格を P_t としたとき，その対数差分は

$$\ln(P_t) - \ln(P_{t-1}) \tag{1}$$

となる．ただし $\ln(x)$ は，x の自然対数を指す．

また，金融系の時系列データには，原系列データの特徴を表すデータの一つに収益率というものが存在する．収益率はいわば現時点での価格が直前の価格から変動した割合である．t 時点での価格データを P_t としたとき，t 時点での収益率 R_t は，

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} \tag{2}$$

で算出される．

また，$\ln(x)$ に対し，x が十分小さい場合は，

$$\ln(1+x) \approx x \tag{3}$$

という近似式が成り立つ．

ここで，式 (3) より，

$$R_t \approx \ln(1+R_t) \tag{4}$$

とし，式 (2) を式 (4) に代入すると，式 (4) は，

$$\begin{aligned} R_t &= \ln(1 + \frac{P_t - P_{t-1}}{P_{t-1}}) \\ &= \ln(\frac{P_{t-1}}{P_{t-1}} + \frac{P_t - P_{t-1}}{P_{t-1}}) \\ &= \ln(\frac{P_t}{P_{t-1}}) \end{aligned} \tag{5}$$

と変形できる．

また，

$$\ln(\frac{a}{b}) = \ln(a) - \ln(b) \qquad (a, b \neq 0) \tag{6}$$

より，式 (5) は

$$R_t = \ln(P_t) - \ln(P_{t-1}) \tag{7}$$

と変形できる．したがって，金融系の時系列データの収益率は原系列データの対数差分と置くことができる．また，原系列データの対数差分は金融商品の価格の変動の割合とみなすことができる．

以降，本研究においては単に収益率と記載した場合は，原系列データの対数差分のことを指し，対数差収益率と表す．

6.2 定常性の調査

時系列データは，全ての時点のデータはそれぞれ一度しか観測することができない．したがって，時系列解析では t 時点での得られた時系列データ x_t を，確率変数 X_t によって算出された実現値とみなし，その確率変数の確率分布に関して何かしらの性質や構造を仮定する．

ここで，解析対象の時系列データが定常性を有していない場合，確率分布の構造の推定が非常に困難である場合が多いため，解析対象のデータの性質を表しかつ定常性を有するデータに変換する必要がある．定常性には弱定常性・強定常性の 2 種類あり，金融・経済データの解析に関しては弱定常性を有している必要がある．

なお，弱定常性を有する時系列データとは，t 時点でのデータ値を算出する確率分布の平均・自己共分散が時間 t に依存せず一定であることを示す．以降，本研究においては単に定常性と記載した場合は弱定常性のことを指す．自己共分散とは，時系列データにおいて t 時点でのデータ値 X_t と過去 $t-j$ 時点でのデータ値 X_{t-j} の値との関係を示した指標のことである．t 時点における平均値を $E(X_t) = \mu_t$ としたとき，X_t と X_{t-j} の自己共分散 $Cov(X_t, X_{t-j})$ は，

$$Cov(X_t, X_{t-j}) = E((X_t - \mu_t)(X_{t-j} - \mu_{t-j})) \tag{8}$$

となる．[10]

時系列データが定常性を有しているか否かを調査するには，対象の時系列データに対して単位根検定 (拡張 Dickey-Fuller 検定 [11][12]) を行う．単位根検定とは対象の時系列データに単位根が存在するか否かを検定するものであり，結果として出力される p 値によって判断する．拡張 Dickey-Fuller 検定の帰無仮説は「データに単位根が存在する.」である．

経済系の時系列データでは有意水準を 10% とする場合が多く，本研究においても有意水準を 10% とする．したがって p 値が 0.1 以下ならばそのデータに単位根は存在しないものとみなす．データに単位根が存在する場合はその時系列データは定常性を有しておらず (定常過程ではない)，存在しない場合はその時系列データは定常性を有する (定常過程である) と言える．

データに単位根が存在する時系列データ (単位根過程) とは，ここでは t 時点でのデータ値 X_t が，ホワイトノイズ ε_t の和分であるとみなせる時系列データのことを指す．なおホワイトノイズとは，平均 $\mu = 0$，分散 σ^2 が一定，自己共分散 $Cov(\varepsilon_t, \varepsilon_{t-j}), (j = 1, 2, ..., t-1)$ が全て 0 である定常過程の時系列データを指す．[14] つまり，単位根過程とは t 時点でのデータ値 X_t を

$$\begin{aligned} X_t &= \sum_{i=1}^{t-1} \varepsilon_i \\ &= \sum_{i=1}^{t-1} X_{t-i} + \varepsilon_i \end{aligned} \tag{9}$$

と表現できるデータ列のことである．

以下に，単位根過程の時系列データの例及びホワイトノイズの例を示す．

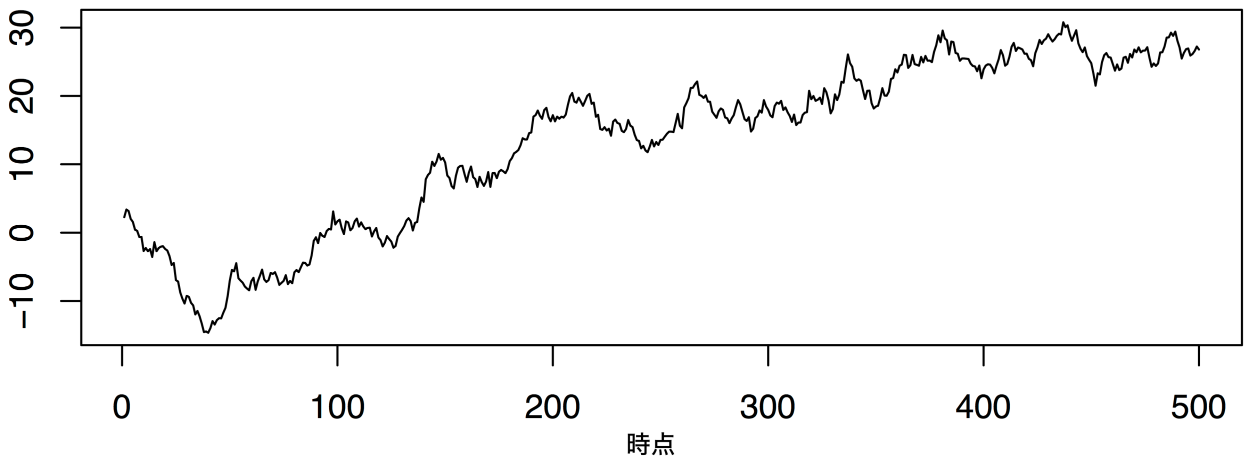


図 1　単位根過程の時系列データ例 1

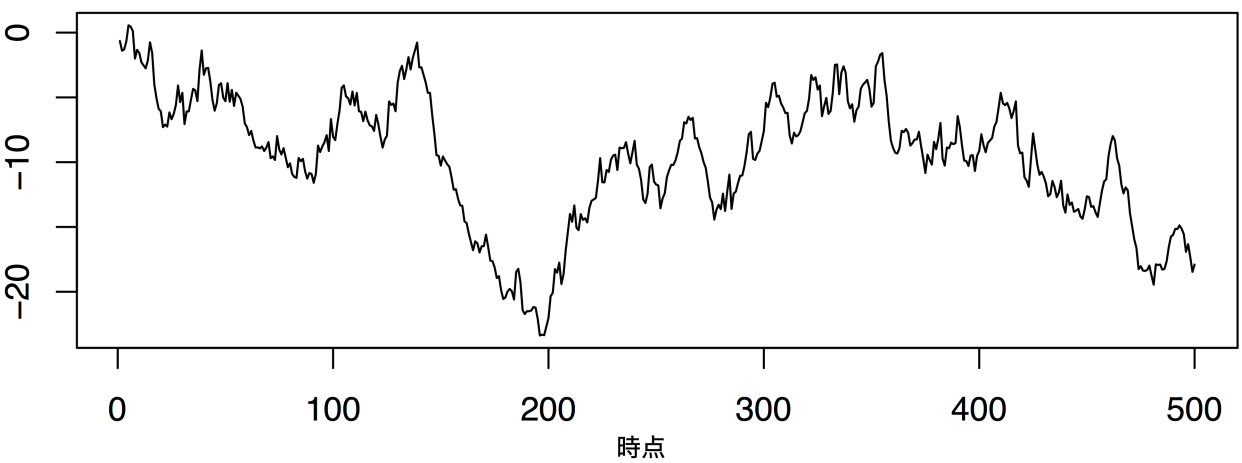


図 2　単位根過程の時系列データ例 2

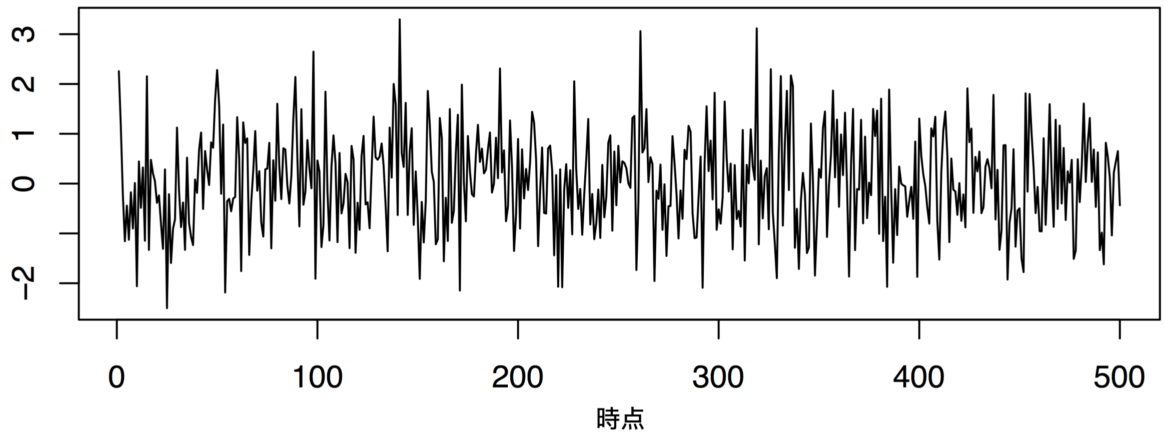


図 3　ホワイトノイズの例 1

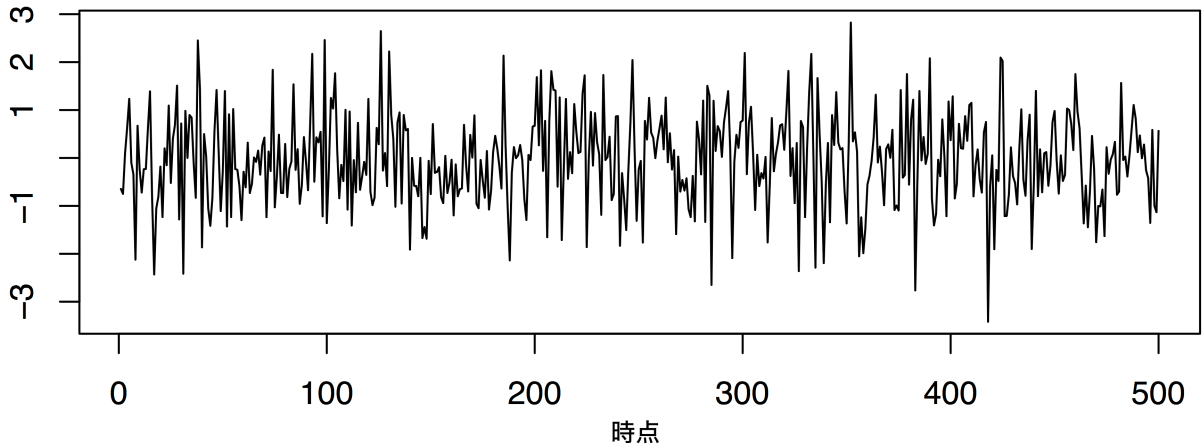


図 4　ホワイトノイズの例 2

ちなみに，単位根過程図 1 及び図 2 は，それぞれホワイトノイズ図 3 及び図 4 の和分である．

6.3 自己相関

時系列データの統計的性質を調査する方法の 1 つに，そのデータの対数差分に対する自己相関を調査することが挙げられる．

自己相関とは，自己共分散同様に t 時点でのデータ値 X_t と過去 $t-j$ 時点でのデータ値 X_{t-j} の値との関係を示した指標のことであり，自己共分散を $t, t-j$ 時点での標準偏差の積で正規化したものである．したがって，$Cov(X_t, X_{t-j})$ を X_t と X_{t-j} の自己共分散，$Var(X_t)$ を時系列データ $X_1, X_2, ..., X_t$ までの分散とすると，t 時点でのデータ値 X_t と過去 $t-j$ 時点でのデータ値 X_{t-j} の自己相関係数は

$$\frac{Cov(X_t, X_{t-j})}{\sqrt{Var(X_t)Var(X_{t-j})}} \tag{10}$$

となる．式 10 で算出された自己相関係数は，0...1 の値を取り，その値が 1 に近いほど $t-j$ 時点でのデータ値 X_{t-j} は，t 時点でのデータ値 X_t に関係があると見なされる．

自己相関を調査する場合には，コレログラムを導出すると良い．コレログラムとは時系列データに対するラグごとの自己相関係数を示したグラフのことである．ラグとは，自己相関係数を算出する上で使用する過去のデータに対する元データからずらした時点数のことで，t 時点と $t-j$ 時点の値の自己相関係数を算出する場合は，そのラグは j にあたる．

6.4 自己回帰モデル

時系列解析では，自己相関の調査後には，AR(n) モデルへの当てはめを試みることがよくされている．AR は Auto Regressive の略であり，自己回帰を意味する．AR(n) モデルとは現時点のデータの値 X_t を過去 n つのデータに係数 $\phi_i (0 \leq i \leq n, i \in Z, 0 < \phi_i < 1)$ をかけて表現する時系列モデルである．n は AR モデルの次数とも言う．すなわち，現時点のデータの値 X_t を，以下の式

$$X_t = c + \sum_{i=1}^{n} \phi_i X_{t-i} + \varepsilon_t \tag{11}$$

で算出する．ただし c は定数項であり，対象のデータ X の総数を m とし，以下の式

$$c = (1 - \phi_1)\frac{1}{m}\sum_{k=1}^{m} X_k \tag{12}$$

で算出される．また，ε_t はホワイトノイズであり，誤差項を示す．

例えば，AR(1) モデルならば，現時点でのデータの値 X_t は，

$$X_t = c + \phi_1 X_{t-1} + \varepsilon_t \tag{13}$$

と表現される．

ただし，AR(n) モデルにおいて，$n = 0$ となった場合は対象のデータを AR(n) モデルを用いて表現することは不可能であるとみなされる．

なお，本研究では，AIC(赤池情報量基準) による AR(n) モデルの次数決定をする．AIC とは，統計モデルの尤もらしさを評価するための指標であり，この指標の値が小さいほど良いモデルとされる．[12][13]

7　BTC 価格の統計的性質の調査結果

7.1　BTC 価格とその収益率

2018 年 10 月 31 日 12 時 00 分 00 秒から 2018 年 11 月 30 日 23 時 55 分 00 秒までの，5 分ごとに観測した BTC の価格のグラフを以下に示す．

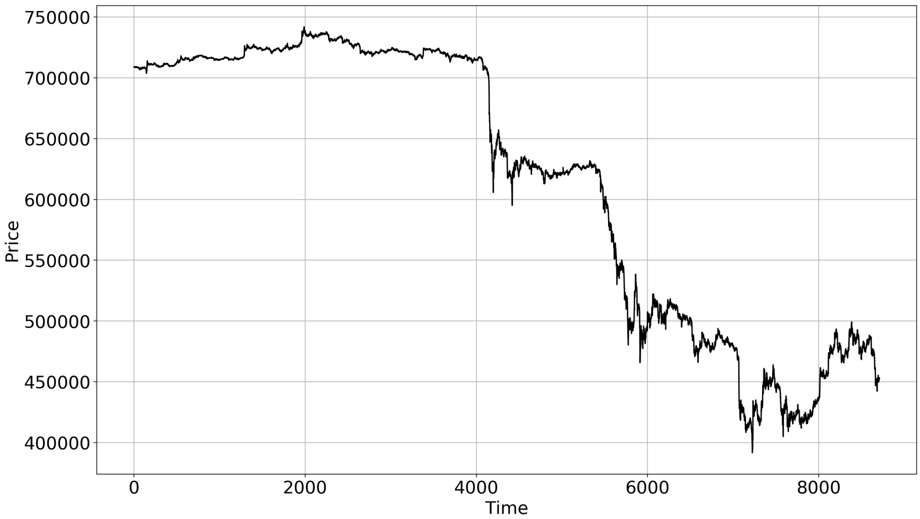


図 5　BTC 価格

また，2018 年 10 月 31 日 12 時 00 分 00 秒からトレンドが転換する直前の 2018 年 11 月 14 日 23 時 55 分 00 秒までの，5 分ごとに観測した BTC の価格のグラフを以下に示す．トレンドとは，時系列データに対する値の大まかな推移傾向を指す．

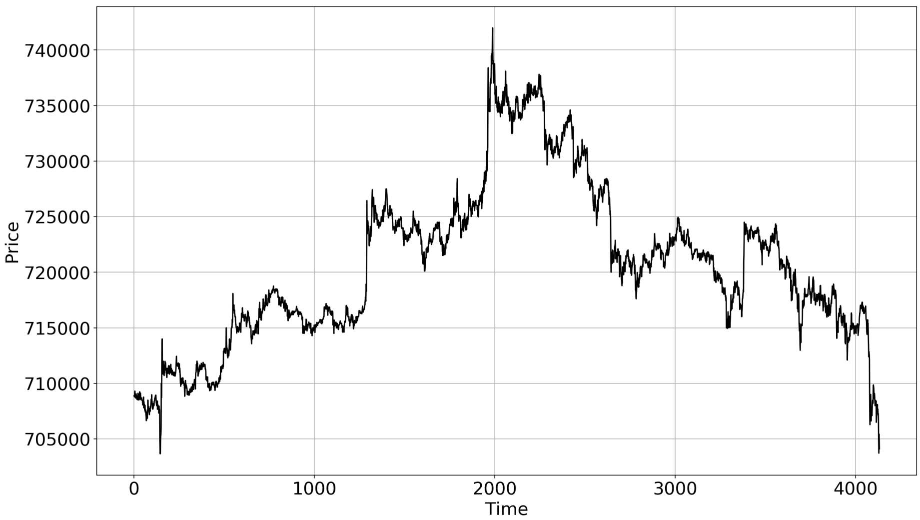


図 6　BTC 価格 2018 年 11 月 14 日 23 時 55 分 00 秒まで

続いて，2018 年 10 月 31 日 12 時 00 分 00 秒から 2018 年 11 月 30 日 23 時 55 分 00 秒までの，5 分ごとに観測した BTC の価格の対数差収益率 (対数差分) のグラフを以下に示す．

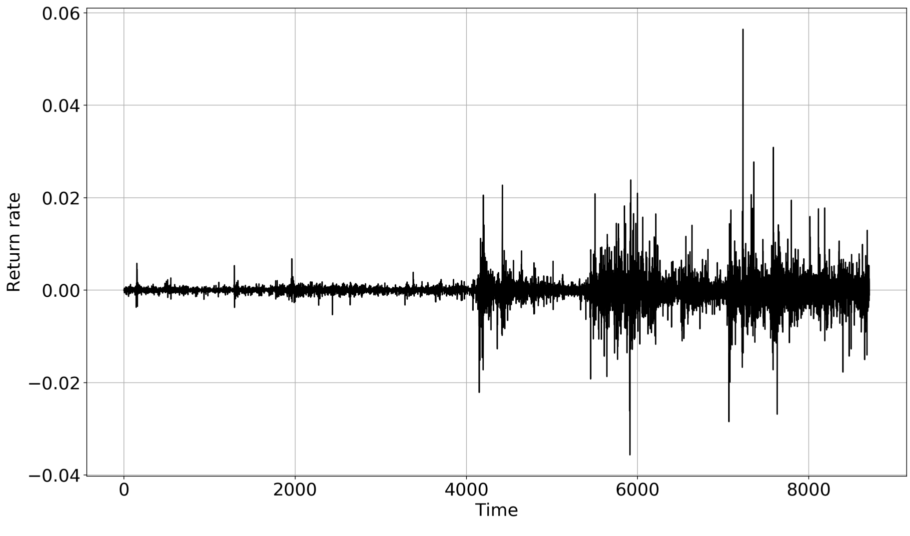


図 7　BTC 価格対数差収益率

また，2018 年 10 月 31 日 12 時 00 分 00 秒からトレンドが転換する直前の 2018 年 11 月 14 日 23 時 55 分 00 秒までの，5 分ごとに観測した BTC の価格の対数差収益率のグラフを以下に示す．

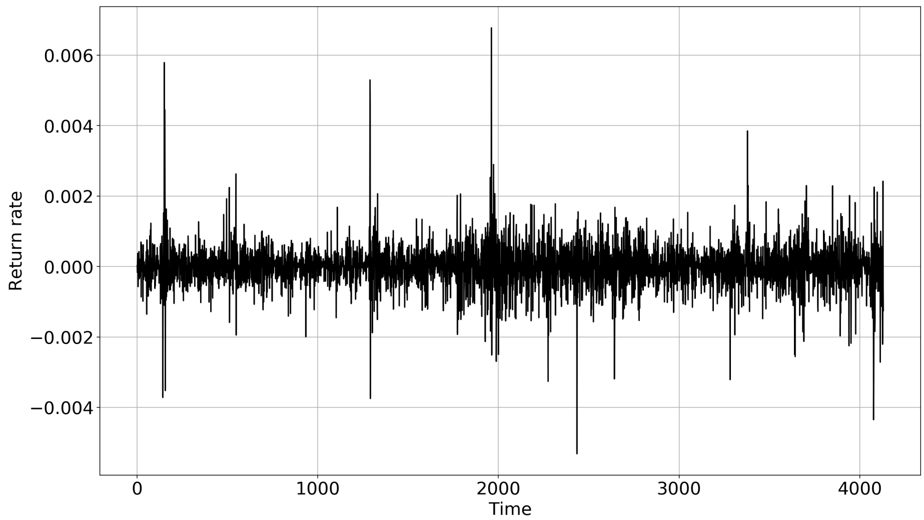


図 8　BTC 価格対数差収益率 2018 年 11 月 14 日 23 時 55 分 00 秒まで

また，トレンドが転換する直前からトレンド転換後，2018 年 11 月 14 日 12 時 00 分 00 秒から 2018 年 11 月 30 日 23 時 55 分 00 秒までの，5 分ごとに観測した BTC の価格の対数差収益率のグラフを以下に示す.

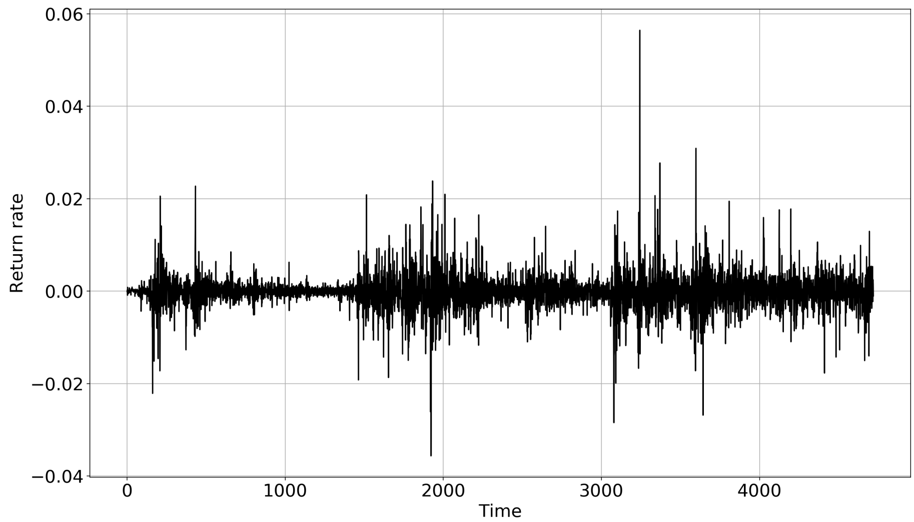


図 9　BTC 価格対数差収益率 2018 年 11 月 14 日 12 時 00 分 00 秒から 2018 年 11 月 30 日 23 時 55 分 00 秒まで

なお，価格及びその収益率を図示する際にトレンド転換前後で分割して図示したが，実際に解析する上ではそのようなことはしていない.

7.2　定常性の調査

まず，2018 年 10 月 31 日 12 時 00 分 00 秒から 2018 年 11 月 30 日 23 時 55 分 00 秒までの，5 分ごとに観測した BTC の価格及びその収益率に対して拡張 Dickey-Fuller 検定を行った結果の p 値を示す．

表 1　単位根検定結果

対象データ	p 値
原系列データ	0.5629
収益率	0.01 未満

単位根検定の結果より，価格データの原系列は定常性を有しておらず，価格データの収益率は定常過程であると言える．

したがって，以降データの性質を調査する際には BTC 価格の収益率を使用する．

7.3　収益率の自己相関の調査結果

BTC の収益率に対して，R による自己相関係数の解析結果のログとコレログラムを以下に示す．なお，R で自己相関係数を解析する場合には，acf() 関数を用いる．このとき解析対象のデータが格納された変数を引数にする．

```
Autocorrelations of series ‘BTC_r’, by lag

     0      1      2      3      4      5      6      7      8      9     10
 1.000  0.030 -0.013  0.043 -0.005 -0.025 -0.013  0.024  0.000 -0.012 -0.030
    11     12     13     14     15     16     17     18     19     20     21
-0.013  0.016 -0.022 -0.017 -0.014  0.008  0.044  0.018 -0.002 -0.028 -0.003
    22     23     24     25     26     27     28     29     30     31     32
-0.022 -0.018  0.007  0.012  0.008 -0.002 -0.005  0.021  0.000  0.023  0.009
    33     34     35     36     37     38     39
 0.012 -0.007  0.019  0.013 -0.010  0.022  0.010
```

図 10　BTC 収益率の自己相関係数

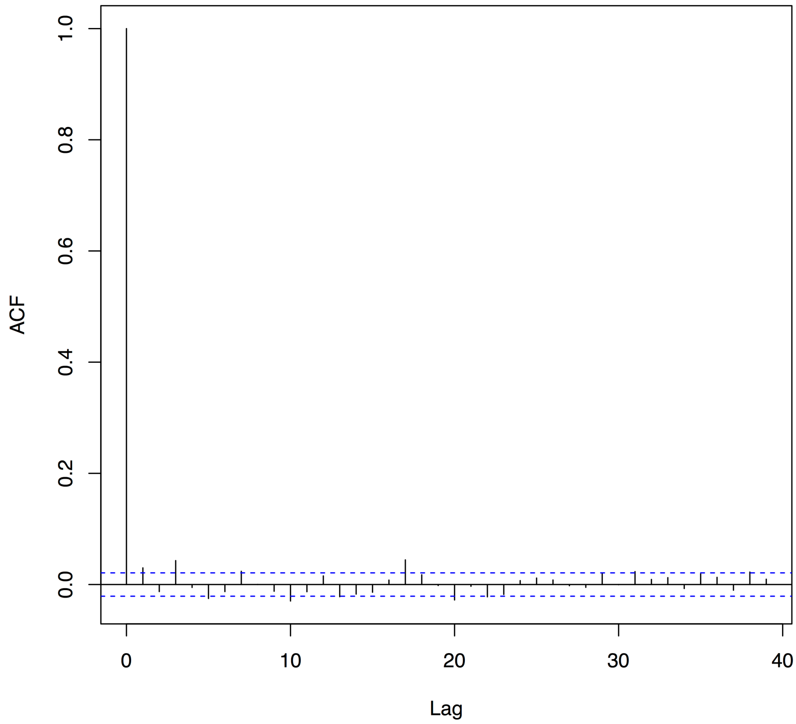


図 11　BTC 収益率のコレログラム

図 10 及び 11 より，BTC 収益率の自己相関係数の絶対値はどの時点においても 0.05 未満であり，有意な自己相関とは言えない．

7.4　収益率に対する自己回帰モデル適用結果

BTC の収益率に対して AIC アルゴリズムを利用し AR(n) モデルの n の値及び係数を求めた結果を示す．R で自己回帰モデルを適用する場合は，ar(x=対象データ, AIC=AIC アルゴリズムを使用するか否か (T/F)) と入力する．

```
Call:
ar(x = BTC_r, aic = T)

Coefficients:
      1        2        3        4        5        6        7        8
 0.0293  -0.0115   0.0476  -0.0087  -0.0260  -0.0137   0.0275   0.0024
      9       10       11       12       13       14       15       16
-0.0121  -0.0332  -0.0119   0.0204  -0.0185  -0.0180  -0.0183   0.0085
     17       18       19       20
 0.0478   0.0159  -0.0047  -0.0328

Order selected 20  sigma^2 estimated as  8.202e-06
```

図 12　R・AIC アルゴリズムによる BTC 収益率に対する AR モデルの係数の解析結果

R で AR モデルの最大次数の指定をせずに次数・係数を求めると，次数 $n = 20$ となり，次数としては大きすぎる結果となった．さらに，全ての係数の絶対値は 0.05 未満であり，有意な係数であるとは考えにくく，AR(0) モデルとみなしても良いような結果となった．

ここで，推定されたモデルの評価を行う．モデル評価に関する情報として，残差の分散，対数尤度，AIC 情報量基準等があり，これらの値によってモデルは評価される．今回は推定された AR(20) モデルと，AR(0) モデルに対してモデル評価に関する情報を比較し，それらの数値に大きな変化が観測されなければ，BTC 価格の収益率は AR(0) モデルとみなす．

以下に，BTC 価格の収益率に対する推定された AR(20) モデル及び AR(0) モデルの評価結果を示す．

```
Call:
arima(x = BTC_r, order = c(20, 0, 0))

Coefficients:
         ar1      ar2     ar3      ar4      ar5      ar6     ar7     ar8
      0.0293  -0.0115  0.0476  -0.0087  -0.0260  -0.0137  0.0275  0.0025
s.e.  0.0107   0.0107  0.0107   0.0107   0.0107   0.0107  0.0107  0.0107
         ar9     ar10     ar11    ar12     ar13     ar14     ar15    ar16
      -0.0121  -0.0332  -0.0119  0.0204  -0.0186  -0.0180  -0.0182  0.0085
s.e.   0.0107   0.0107   0.0107  0.0107   0.0107   0.0107   0.0107  0.0107
        ar17    ar18     ar19     ar20  intercept
      0.0478  0.0159  -0.0047  -0.0328     -1e-04
s.e.  0.0107  0.0107   0.0107   0.0107      0e+00

sigma^2 estimated as 8.182e-06:  log likelihood = 38653.53,  aic = -77263.05
```

図 13　AR(20) モデルの評価

```
Call:
arima(x = BTC_r, order = c(0, 0, 0))

Coefficients:
      intercept
         -1e-04
s.e.      0e+00

sigma^2 estimated as 8.269e-06:  log likelihood = 38607.35,  aic = -77210.7
```

図 14　AR(0) モデルの評価

図 13 及び 14 より，AR(20) 及び AR(0) の残差の分散，対数尤度，AIC 情報量基準を以下の表に示す．

表 2　AR(20) 及び AR(0) モデルの評価結果

評価項目	AR(20)	AR(0)
残差の分散	8.182×10^{-6}	8.269×10^{-6}
対数尤度	38653.53	38607.35
AIC 情報基準量	-77263.05	-77210.7

表 2 より，残差の分散，対数尤度，AIC 情報量基準は両者ともさほど変わらない．従って，本研究では BTC の価格の収益率は AR(0) モデルとみなす．

8 価格データ解析及び投資戦略の検討

8.1 方針

BTC 価格の収益率に対して統計的性質を調査した結果，定常過程ではあるが自己相関はなく，AR モデルでの表現も不適であるとみなされた．このようなデータは時系列解析においては「完全にランダムに発生している」データとみなされるため，確率分布の性質や構造の推定は難しい．

従って，価格データそのものを統計的に解析し投資戦略を検討する．価格データそのものは単位根過程といえるため，解析手法は限られてくるが，本研究では「テクニカル解析」を利用し解析を行う．

テクニカル解析とは，金融系時系列データにおいて将来の価格の変動を過去に発生した価格や出来高等をもとに予測する統計的手法である．テクニカル解析の手法は数多く存在するが，本研究では

- 単純移動平均線
- MACD
- 単純移動平均乖離率

を利用する．

8.2 単純移動平均線

単純移動平均線とは，移動平均線の一種であり，時系列データに対して現時点から過去一定期間のデータの平均をとり，その平均値を結んだ線のことである．ある時系列データの実現値を $X_1, X_2, ...$，単純移動平均線の期間を p としたとき，任意の時点 t での単純移動平均の値 $SMA(t,p)$ は，

$$SMA(t,p) = \frac{1}{p}\sum_{i=0}^{p-1} X_{t-i} \tag{14}$$

と算出される．以下に，期間 5 時点の単純移動平均の算出例を示す．

表 3　単純移動平均算出例

時点	1	2	3	4	5	6	7	8	9	10
実現値	10	20	30	40	50	20	40	60	80	100
移動平均	-	-	-	-	30	32	36	42	50	60

移動平均線を利用した買い時 (買いシグナル)，売り時 (売りシグナル) の目安の一つにゴールデンクロス (GC)，デッドクロス (DC) がある．ゴールデンクロスとは，原系列データに対する期間の異なる 2 つの移動平均線をチャート上に表示し，短期の移動平均線が長期の移動平均線を下から上に突き抜けたことを指す．反対にデッドクロスとは，短期の移動平均線が長期の移動平均線を上から下に突き抜けたことを指す．実際のデータによる単純移動平均線のゴールデンクロス・デッドクロスの例を示す．

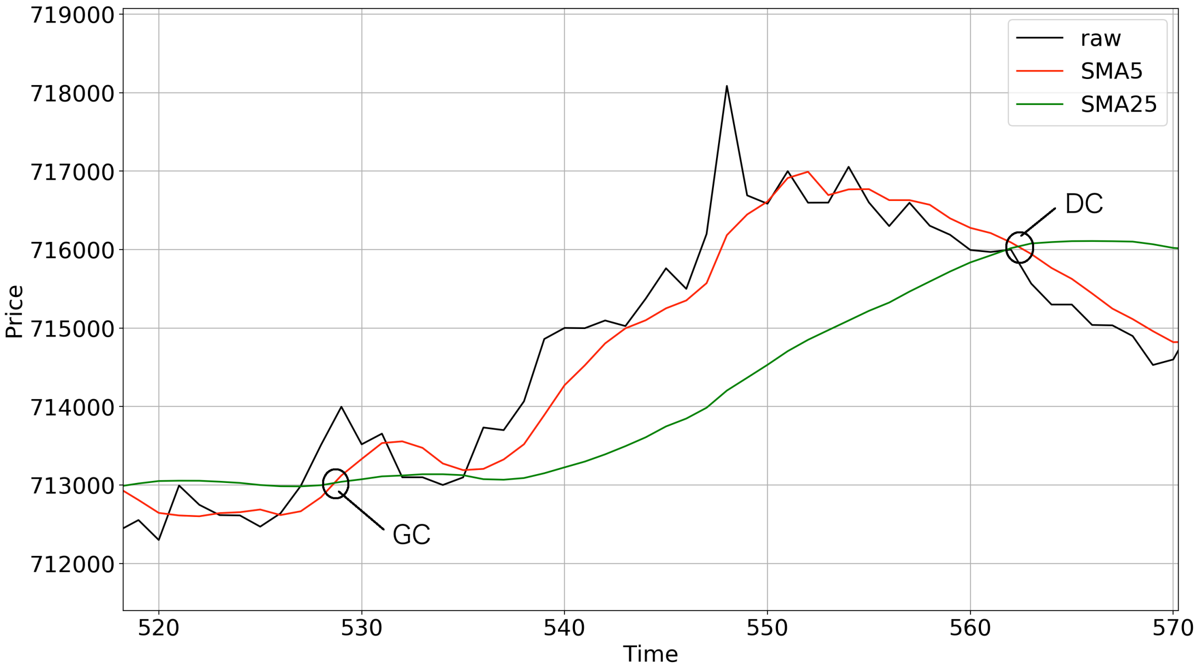


図 15　単純移動平均線 GC・DC の一例

図 15 中における GC の時点で購入し，DC で売却した場合，利益をあげることができる．なお，図 15 における凡例の raw は原系列データ，SMA5 は 5 時点単純移動平均線，SMA25 は 25 時点単純移動平均線である．

本研究では，期間 5，25，75 時点の 3 つの単純移動平均を算出し，5・25 時点，25・75 時点の組み合わせを検証する．5・25 時点の組み合わせは，短時間での売買によって利益をあげることが目的の場合に使用される．また，25・75 時点では，小さな上昇・下降トレンドにそって売買し利益をあげることが目的の場合に使用される．

8.3　MACD

MACD(Moving Average Convergence Divergence) は，移動平均収束拡散法とも言われている，2 つの異なる期間の指数平滑移動平均線の差をとった MACD ラインと，その MACD ラインの任意の期間の単純移動平均線である MACD シグナルから構成されているテクニカル指標である．[15]

ある時系列データの実現値を $X_1, X_2, ...$，指数平滑移動平均線の期間を p としたとき，任意の時点 t での指数平滑移動平均の値 $EMA(t, p)$ は，

$$EMA(t,p) = EMA(t-1,p) + \frac{2}{p+1}(X_t - EMA(t-1,p)), \qquad (t > p) \tag{15}$$

と算出される．なお，$\frac{2}{p+1}$ は，平滑化定数と呼ばれ，α と表される場合もある．ただし，p 時点での指数平滑移動平均の値 $EMA(t = p, p)$ は，

$$EMA(p,p) = \frac{1}{p}\sum_{i=0}^{p-1} X_{p-i} \tag{16}$$

として計算される．以下に，期間 5 時点の指数平滑移動平均の算出例を示す．

表 4　指数平滑移動平均算出例

時点	1	2	3	4	5	6	7	8	9	10
実現値	10	20	30	40	50	20	40	60	80	100
移動平均	-	-	-	-	30	$26.\dot{6}$	$31.\dot{1}$	$40.\dot{7}4\dot{0}$	53.827...	69.218...

従って，ある時系列データの実現値を $X_1, X_2, ...$，指数平滑移動平均線の 2 つの期間を $p_1, p_2 (p_1 < p_2)$ としたとき，任意の時点 t での MACD ラインの値 $MACD(t, p_1, p_2)$ は，

$$MACD(t, p_1, p_2) = EMA(t, p_1) - EMA(t, p_2) \tag{17}$$

である．また，期間 p における任意の時点 t での MACD シグナル $MS(t, p)$ は，t 時点での MACD ラインの値を $MACD_t$ とし，

$$MS(t,p) = \frac{1}{p}\sum_{i=0}^{p-1} MACD_{t-i} \tag{18}$$

と表される．

MACD の買いシグナルは，$MACD(t, p_1, p_2) > MS(t, p)$ かつ $MACD(t-1, p_1, p_2) \leq MS(t-1, p)$ となった時点である．また，売りシグナルは，$MACD(t, p_1, p_2) < MS(t, p)$ かつ $MACD(t-1, p_1, p_2) \leq MS(t, p)$ となった時点である．実際のデータによる MACD のゴールデンクロス・デッドクロスの例を示す．

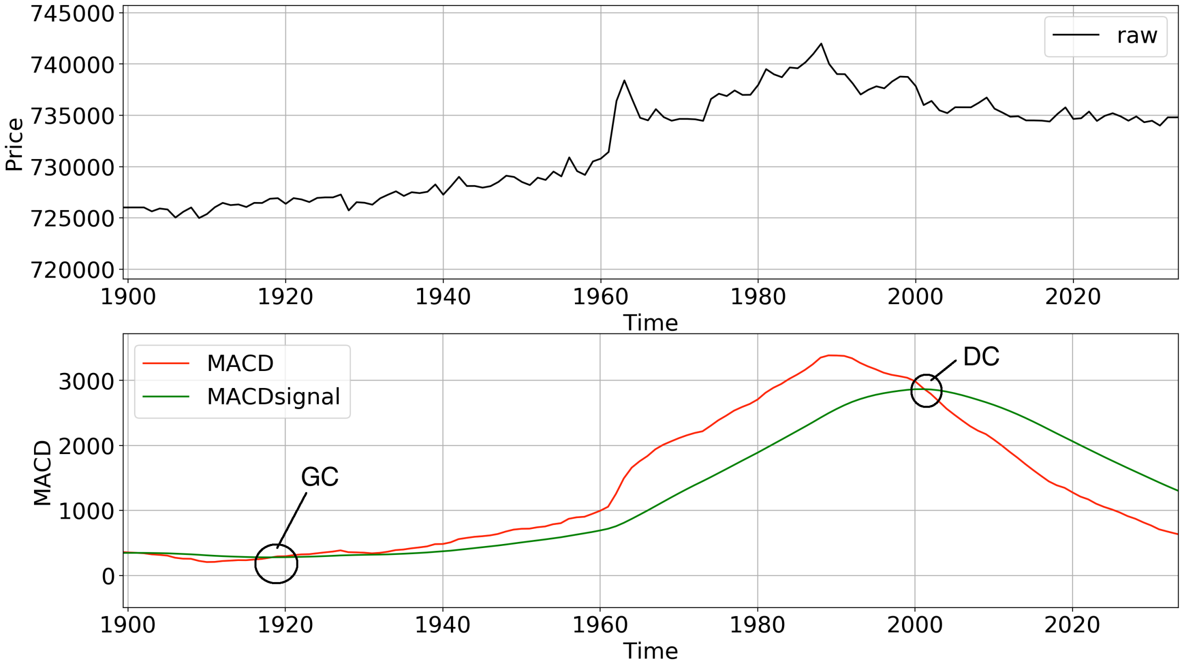


図 16　MACD における GC・DC の一例

なお本研究では，$MACD(t, 12, 26), MS(t, 9)$ の場合を検証する．

8.4　単純移動平均線乖離率

単純移動平均線乖離率とは，価格の単純移動平均線からの離れている割合から買い時・売り時を判断するテクニカル解析手法である．時点 t における価格を P_t，任意の期間 p の単純移動平均線の値を $SMA(t,p)$ とした場合，t 時点の移動平均線乖離率 D_t は，

$$D_t = \frac{P_t - SMA(t,p)}{SMA(t,p)} \tag{19}$$

と算出される．

このテクニカル指標を利用した買い時・売り時に関しては，あらかじめ買い時・売り時とする移動平均線乖離率の値を決めておくことである．例えば，価格が移動平均線から下方に 5%，上方に 5% 乖離した場合に買い時・売り時とする，すなわち移動平均線乖離率の値が -0.05，0.05 となった場合に買い時・売り時とするということである．

本研究では，まず 2018 年 10 月 01 日 00 時 00 分 00 秒から 2018 年 10 月 31 日 23 時 55 分 00 秒の 5 分ごとに観測された価格の，各時点における任意の期間 p の単純移動平均線に対する乖離率を算出し，その乖離率の下位 2.5% の値 d_1，上位 2.5% の値 d_2 を抽出する．そして，実際に扱う価格データに対する任意の期間 p の単純移動平均線乖離率の値が d_1, d_2 になった場合に買い時・売り時とする．本研究では，単純移動平均線乖離率の期間を 25 時点の場合と 100 時点の場合で検証する．

8.5　売買のルール

本研究では，各テクニカル解析の手法の「買い時」に購入し，「売り時」に売却すると仮定し，空売り [16] は行わないものとする．ロスカット等「売り時」となる以前に決済をすることは本研究では行わないものとする．ロスカットとは，ある金融商品を注文し保有している状態で，損失が一定以上の水準に達した場合，その注文に対する決済を行い損失の拡大を防ぐことである．BTC の「買い時」の価格と「売り時」の価格の差分をとり，その差分が正の値ならば「利確」，負の値ならば「損失」とし売買結果とする．このとき損益率 $r_t(t = 1, 2, ...)$ を計算する．損益率は 売却価格 ÷ 購入価格 とする．

最終的に，売買による差分を合計し合計損益額・合計損益率及び勝率を計算する．勝率は，売買回数に対する「利確」とされた回数の割合である．

なお，本研究では売買による手数料・スリッページ (購入価格と売却価格の差異) による損失は無視することが可能であるとする．

9　結果

9.1　単純移動平均線

まず，2018 年 10 月 31 日 12 時 00 分 00 秒から 2018 年 11 月 30 日 23 時 55 分 00 秒における原系列データと 5・25 時点単純移動平均線のチャートを示す．チャートは，はじめに全期間のものを示し，続いて期間を 5 分割したものを順に示す．

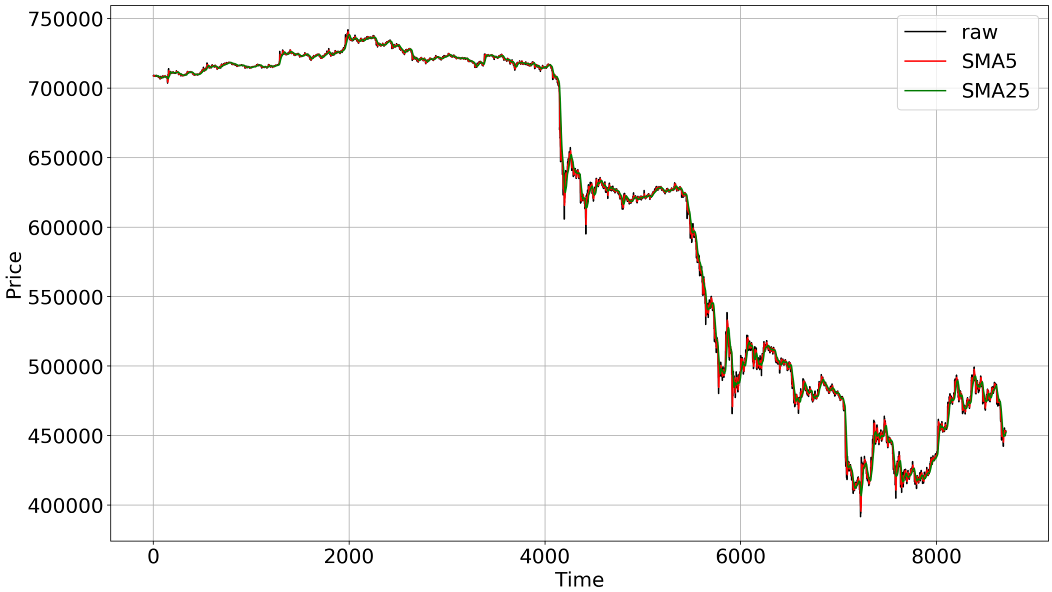


図 17　5・25 時点単純移動平均線チャート全期間

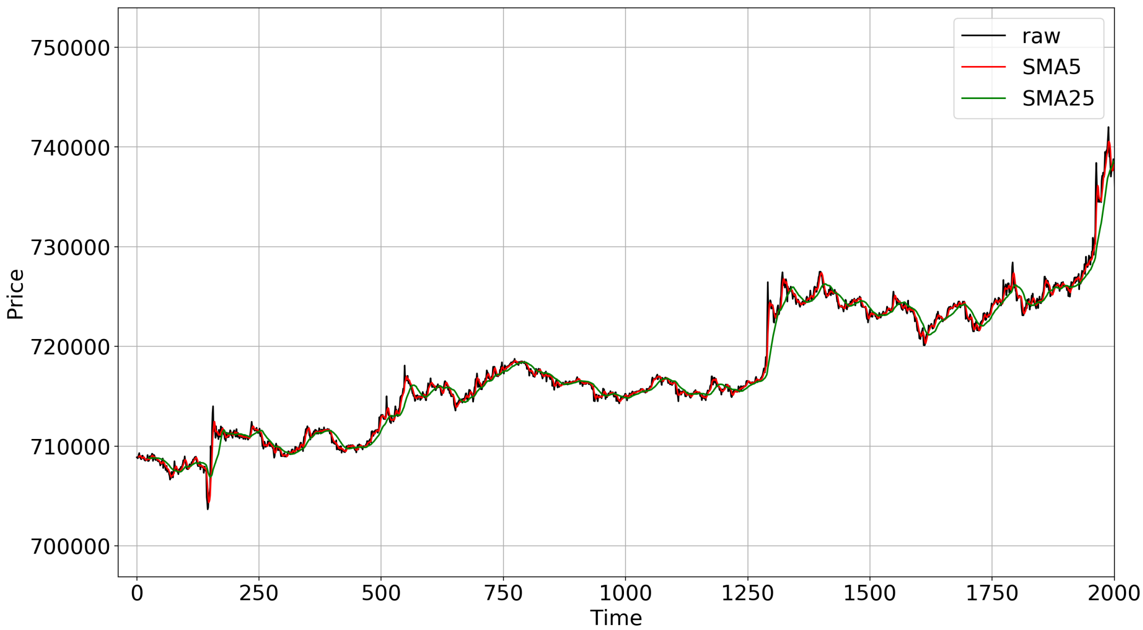


図 18　5・25 時点単純移動平均線チャート 1/5

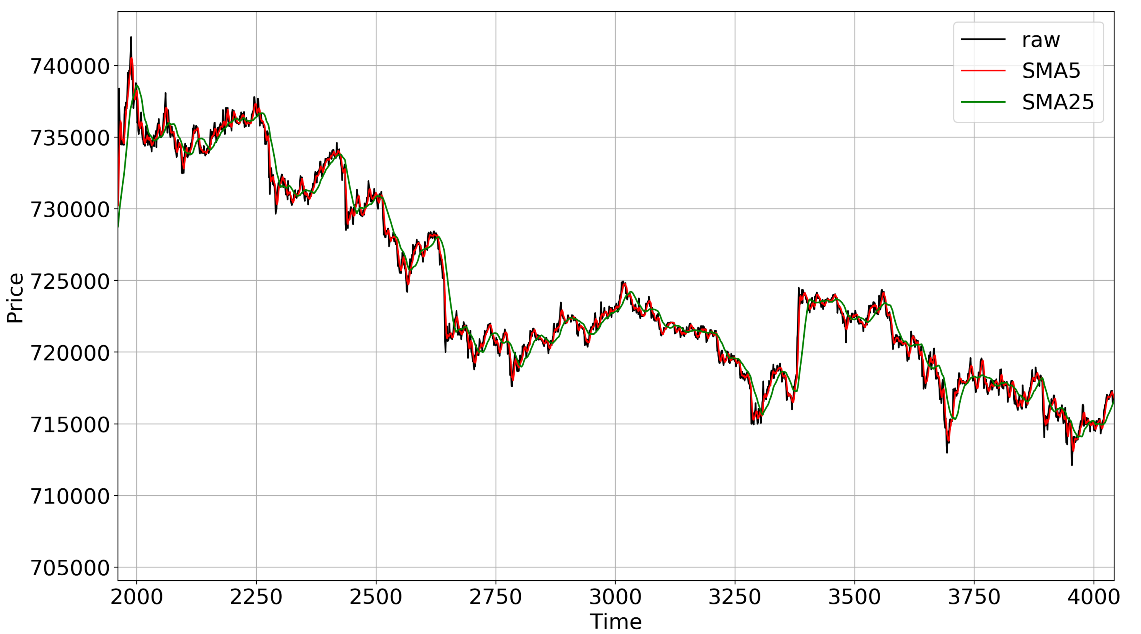


図 19　5・25 時点単純移動平均線チャート 2/5

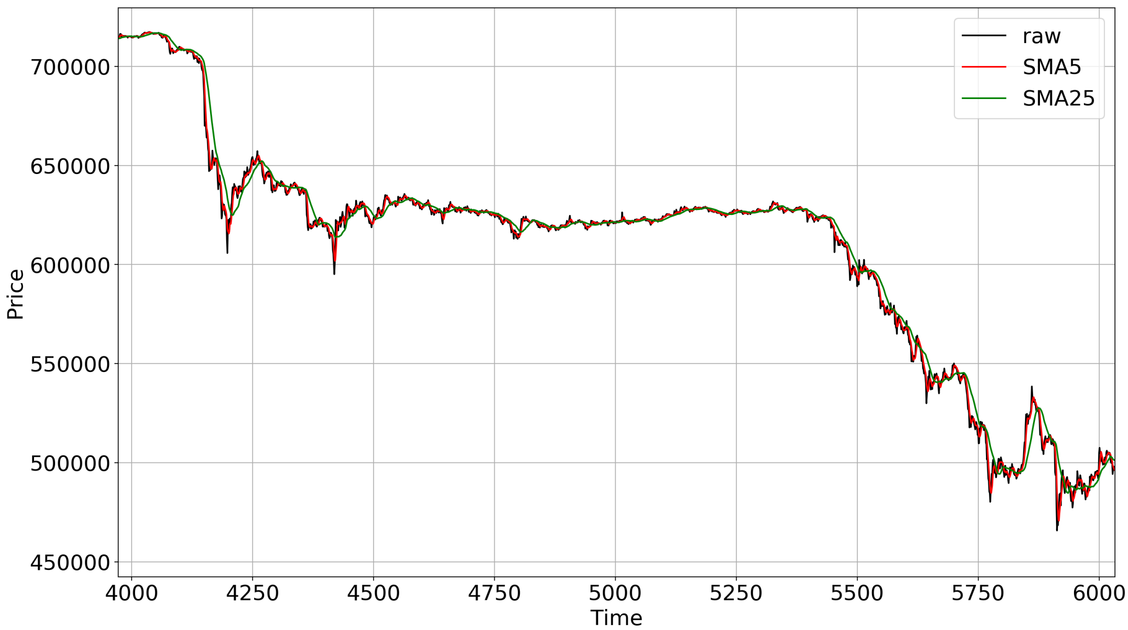


図 20　5・25 時点単純移動平均線チャート 3/5

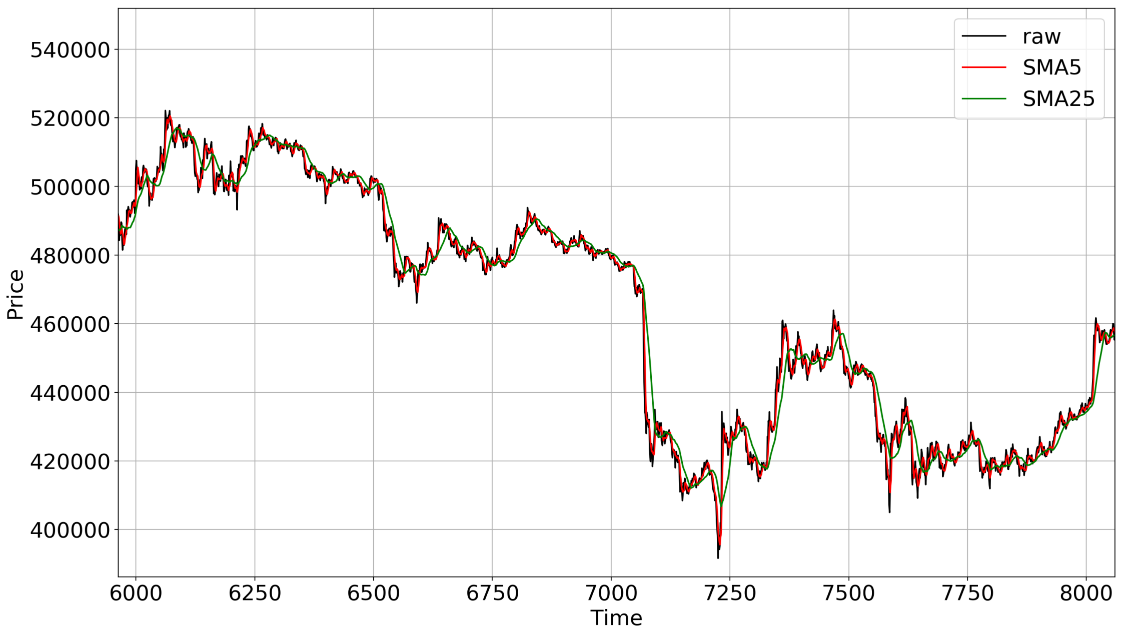


図 21　5・25 時点単純移動平均線チャート 4/5

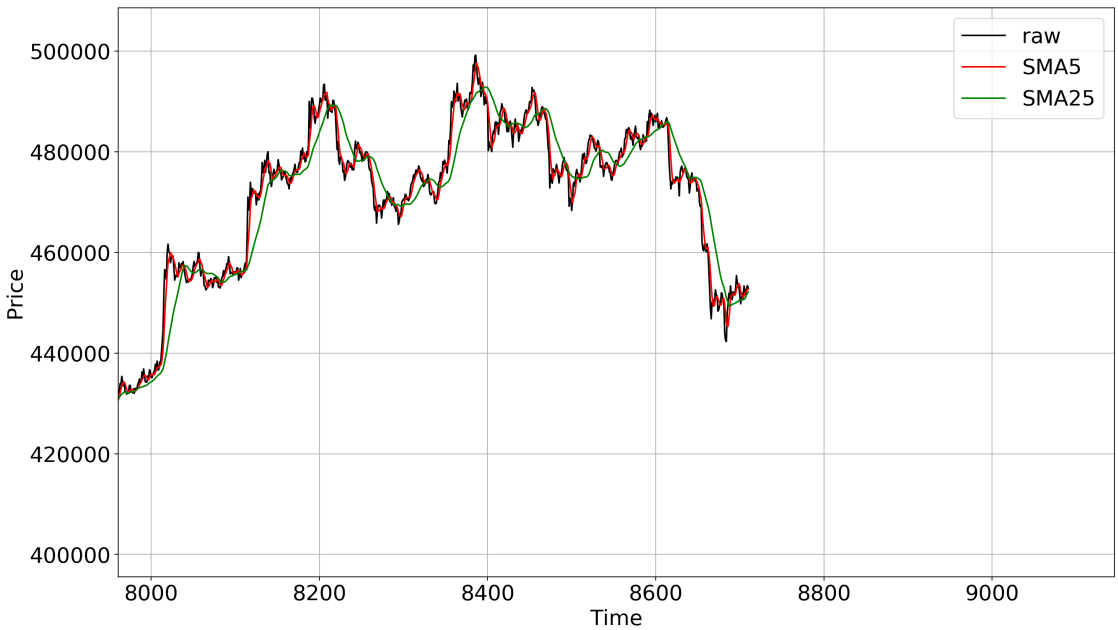
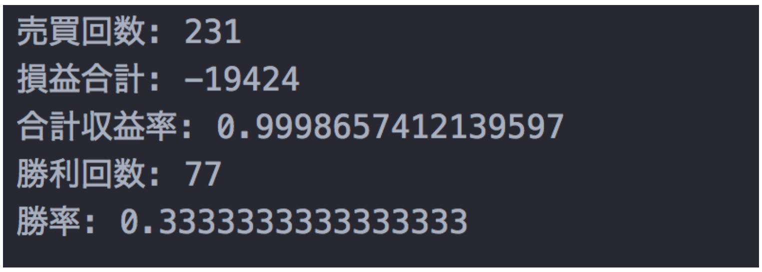


図 22　5・25 時点単純移動平均線チャート 5/5

5・25 時点単純移動平均線における売買の最終結果のログを以下に示す.



```
売買回数: 231
損益合計: -19424
合計収益率: 0.9998657412139597
勝利回数: 77
勝率: 0.3333333333333333
```

図 23　5・25 時点単純移動平均線売買結果

2018 年 10 月 31 日 12 時 00 分 00 秒から 2018 年 11 月 30 日 23 時 55 分 00 秒における原系列データと 25・75 時点単純移動平均線のチャートを示す．

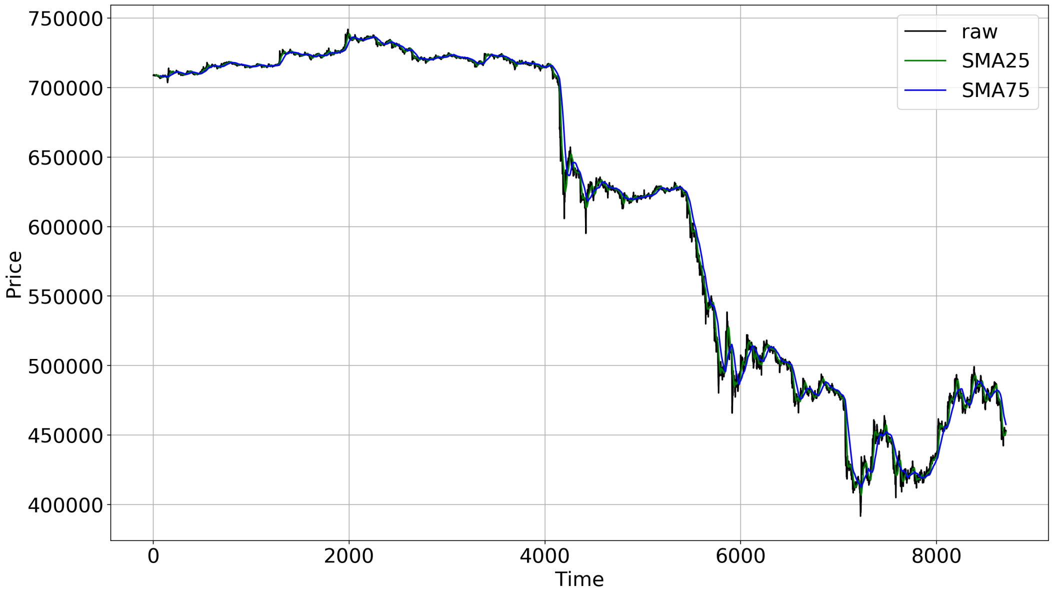


図 24　25・75 時点単純移動平均線チャート全期間

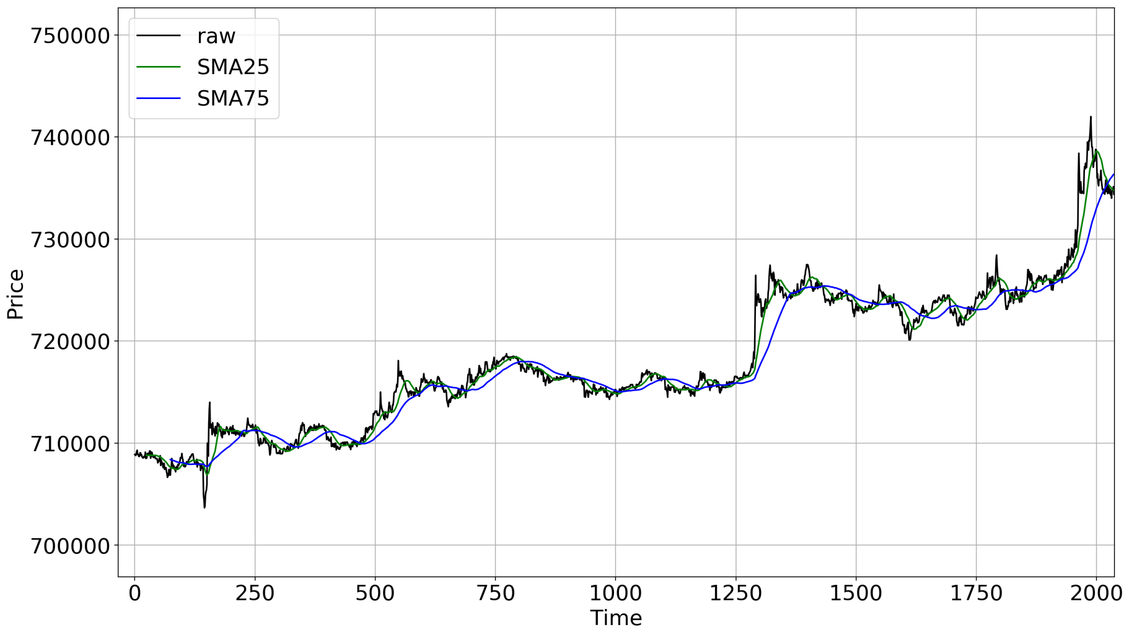


図 25　25・75 時点単純移動平均線チャート 1/5

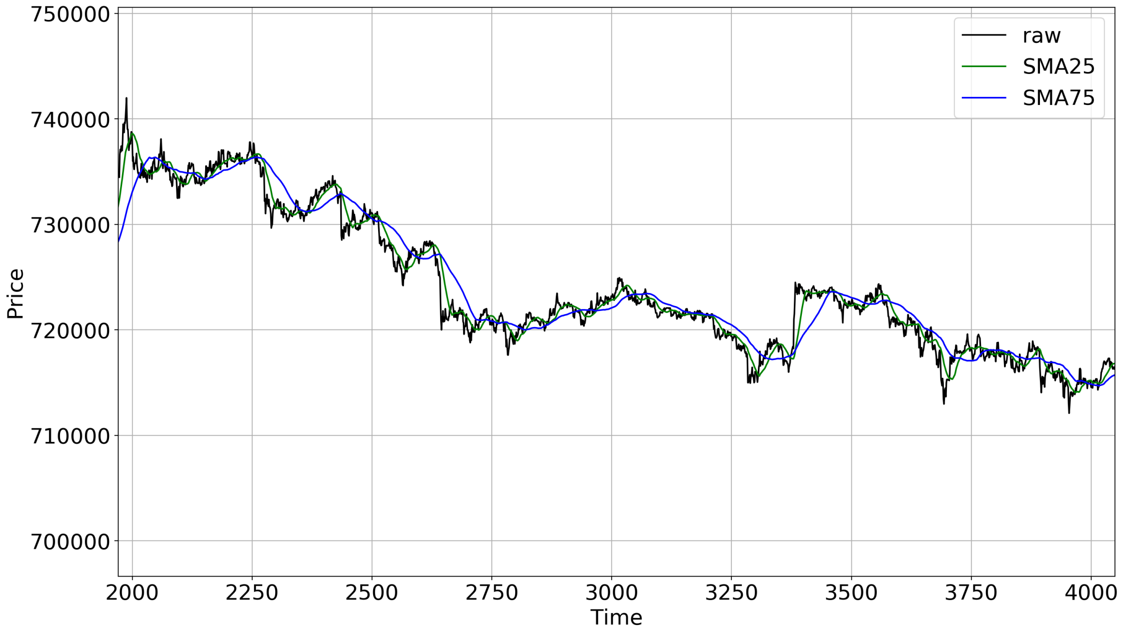


図 26　25・75 時点単純移動平均線チャート 2/5

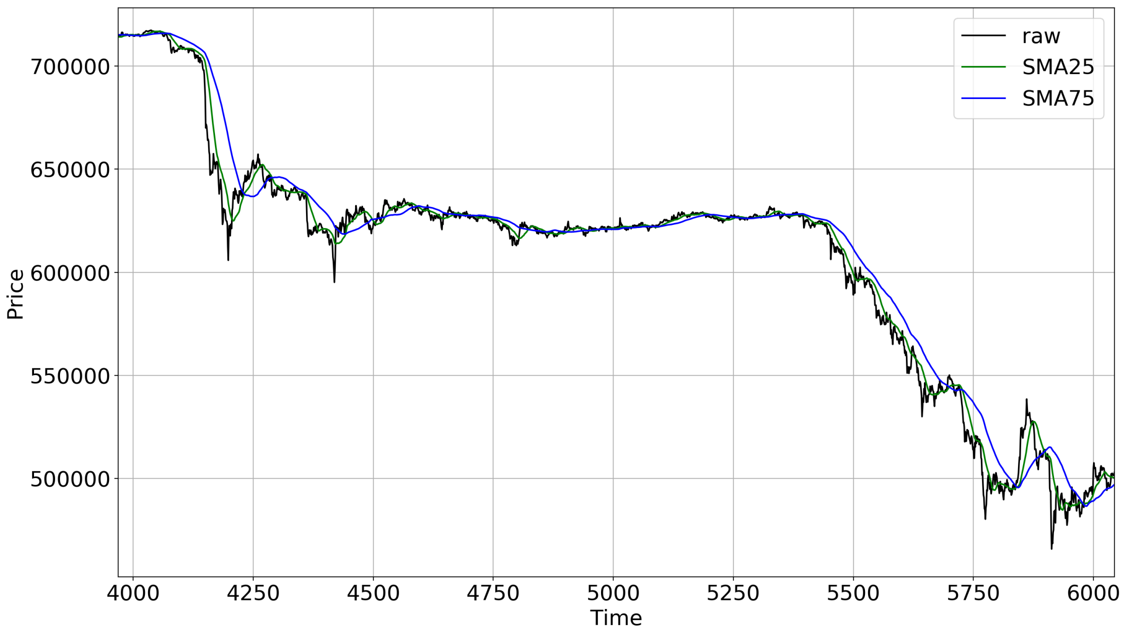


図 27　25・75 時点単純移動平均線チャート 3/5

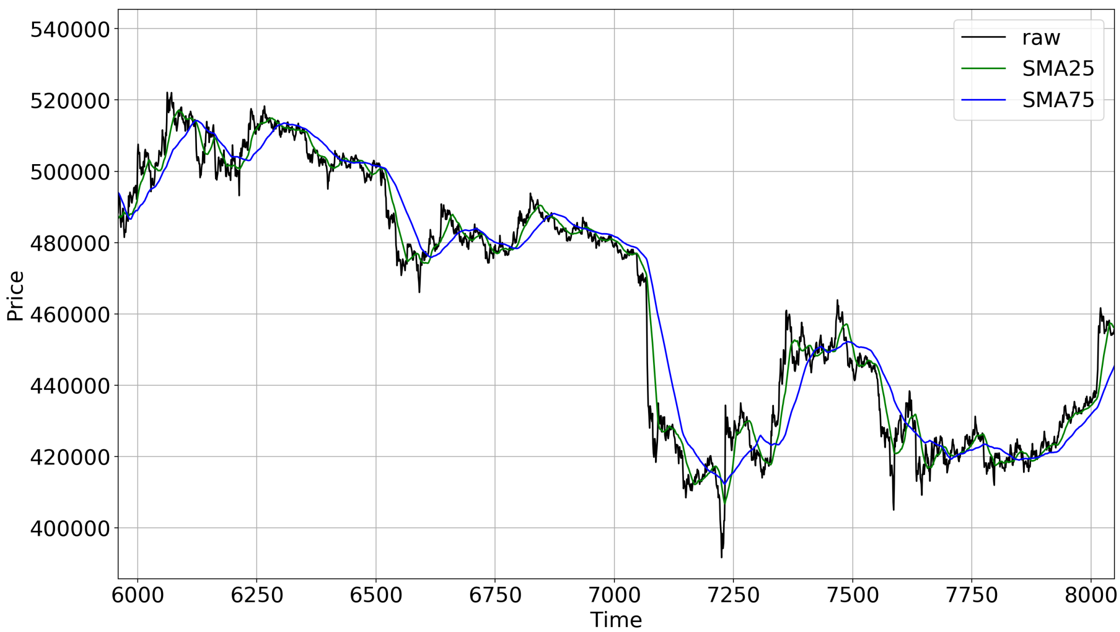


図 28　25・75 時点単純移動平均線チャート 4/5

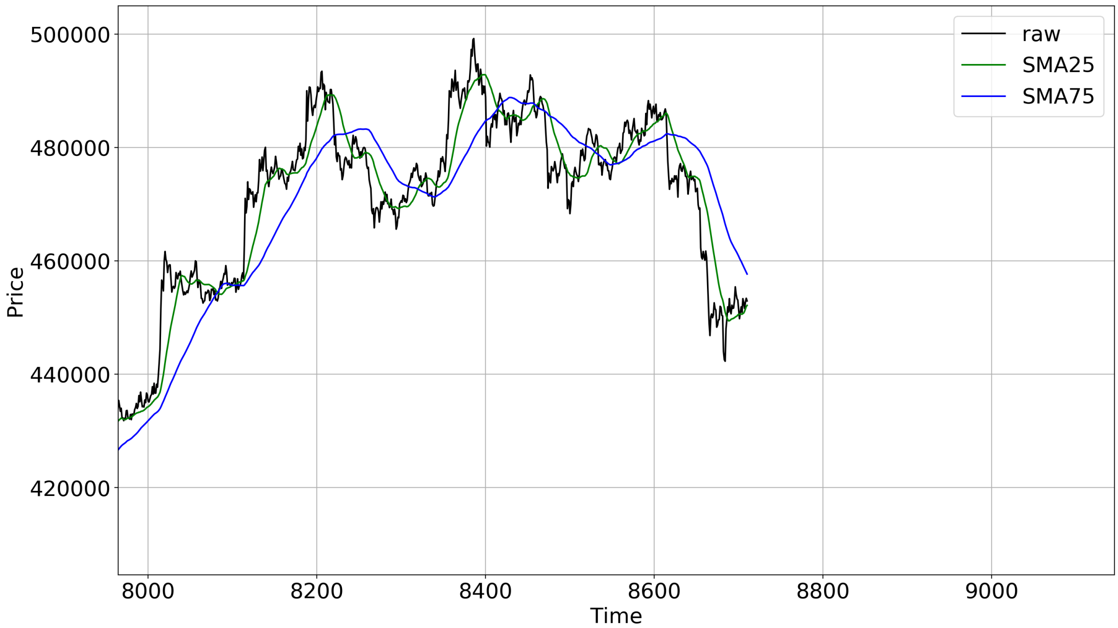


図 29　25・75 時点単純移動平均線チャート 5/5

25・75 時点単純移動平均線における売買の最終結果のログを以下に示す．

```
売買回数: 69
損益合計: 2894
合計収益率: 1.0004002916851853
勝利回数: 20
勝率: 0.2898550724637681
```

図 30　25・75 時点単純移動平均線売買結果

9.2　MACD

2018 年 10 月 31 日 12 時 00 分 00 秒から 2018 年 11 月 30 日 23 時 55 分 00 秒における原系列データと期間 12，26MACD，期間 9MACD シグナルのチャートを示す．チャートは，単純移動平均線同様に，はじめに全期間のものを示し，続いて期間を 5 分割したものを順に示す．

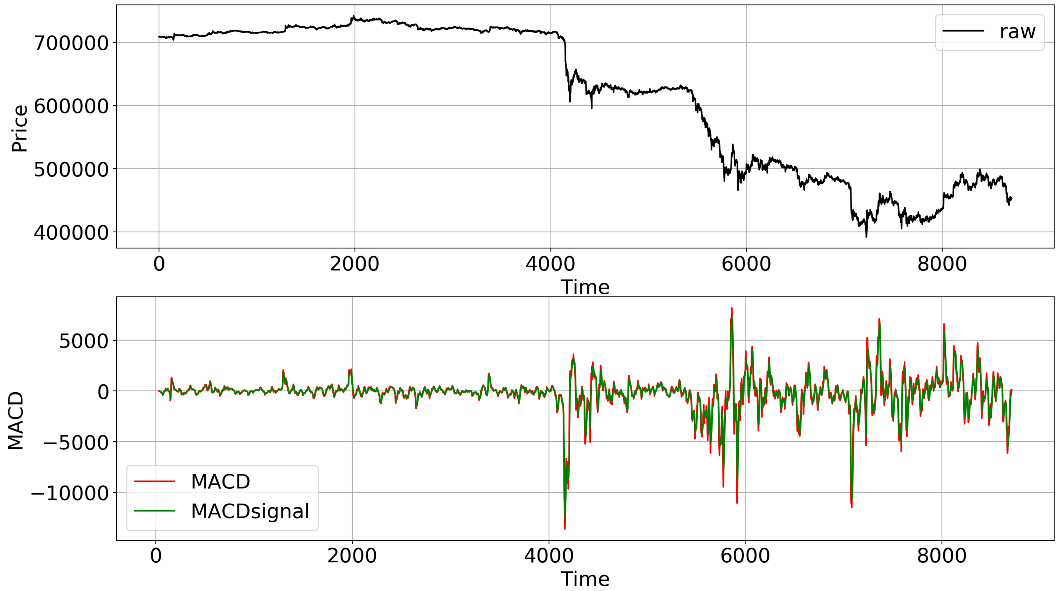


図 31　期間 12・26MACD，期間 9 シグナルチャート全期間

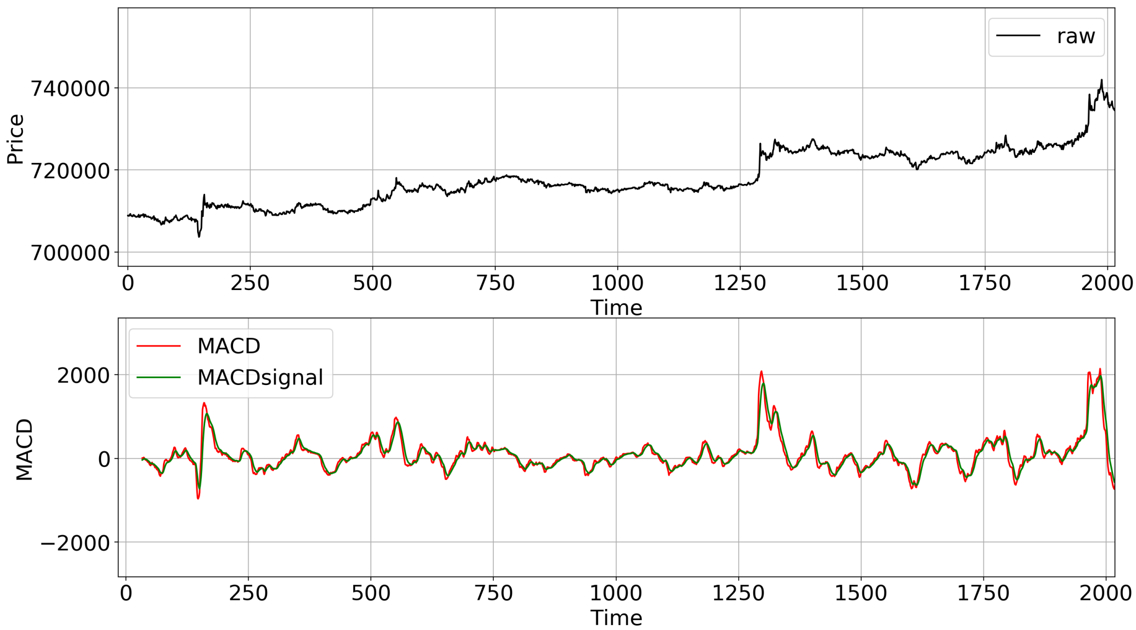


図 32　期間 12・26MACD，期間 9 シグナルチャート 1/5

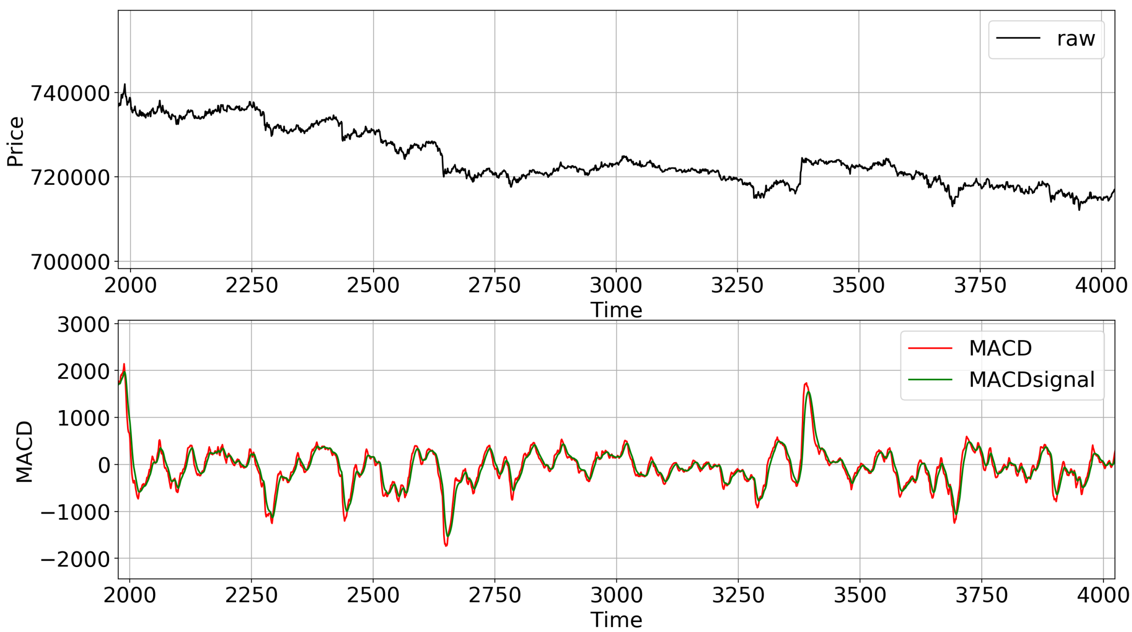


図 33　期間 12・26MACD，期間 9 シグナルチャート 2/5

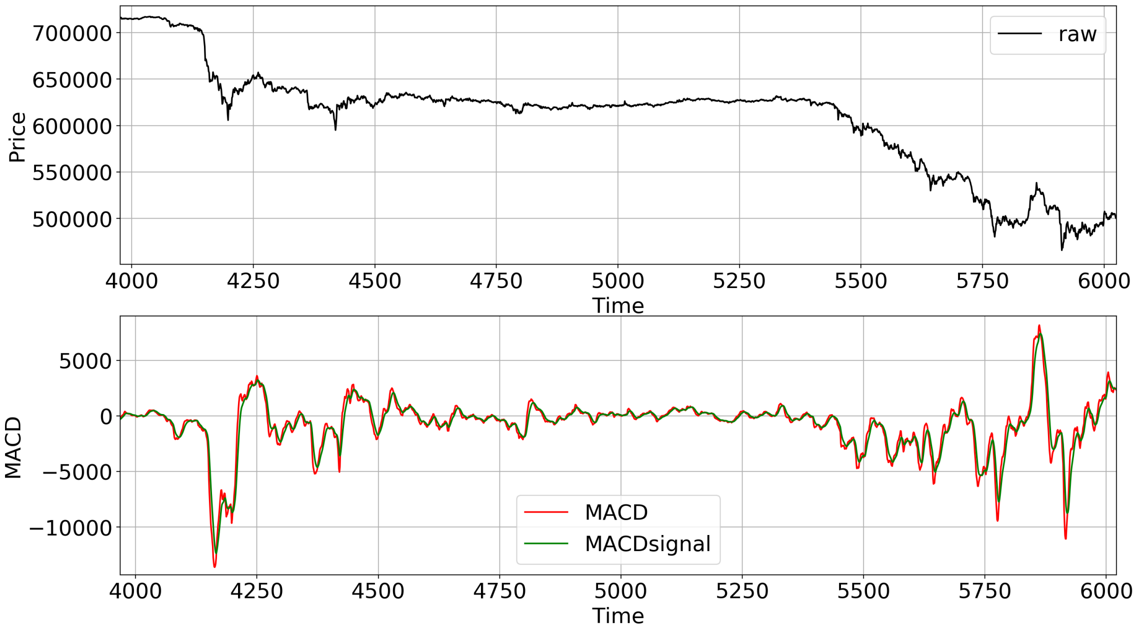


図 34　期間 12・26MACD，期間 9 シグナルチャート 3/5

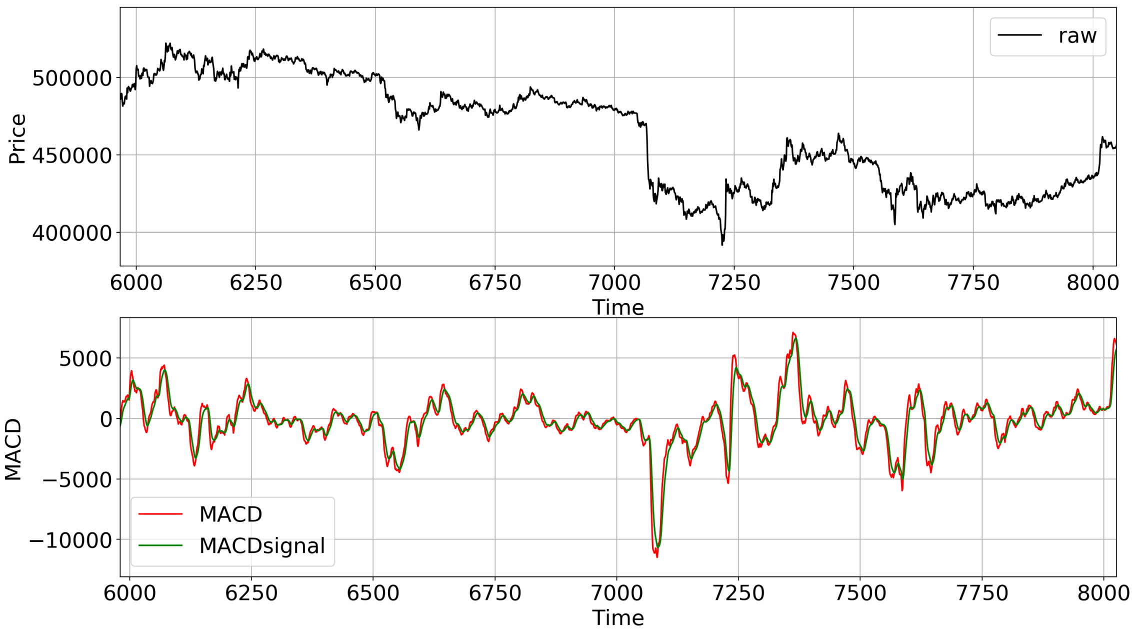


図 35　期間 12・26MACD，期間 9 シグナルチャート 4/5

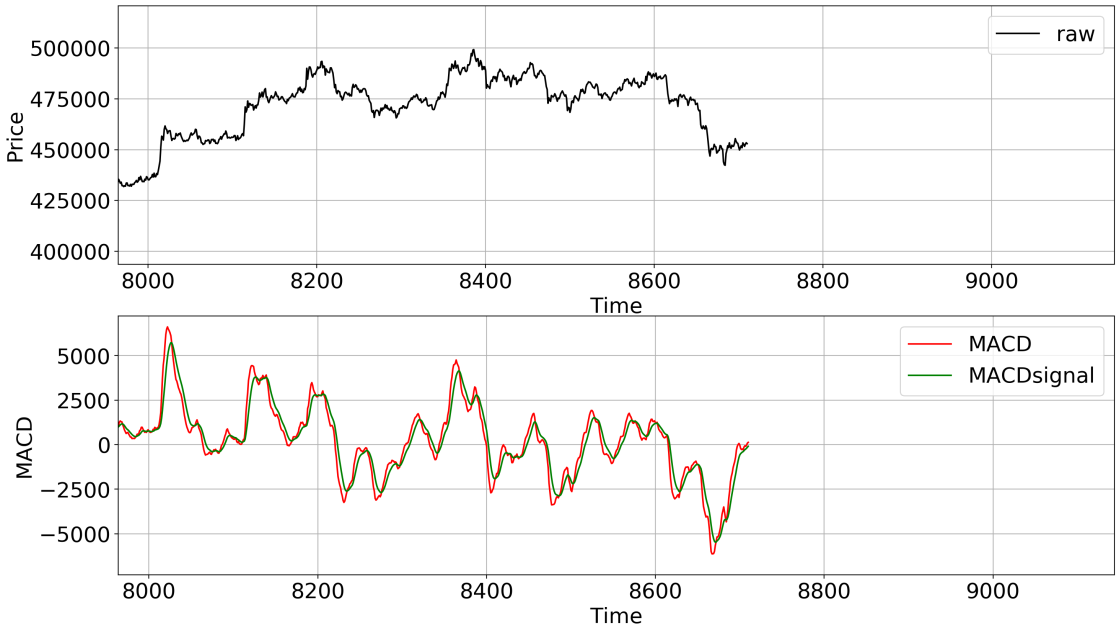


図 36　期間 12・26MACD，期間 9 シグナルチャート 5/5

期間 12・26MACD，期間 9 シグナルにおける売買の最終結果のログを以下に示す．

```
売買回数: 354
損益合計: -153465
合計収益率: 0.9993579385795374
勝利回数: 112
勝率: 0.3163841807909605
```

図 37　期間 12・26MACD，期間 9 シグナル売買結果

9.3　単純移動平均線乖離率

まず，2018 年 10 月 01 日 00 時 00 分 00 秒から 2018 年 10 月 31 日 23 時 55 分 00 秒の 5 分ごとに観測された価格の，各時点における 25 時点単純移動平均線乖離率のヒストグラムを示す．以下に示す乖離率の分布の下位 2.5% の値を「買い時」の乖離率，上位 2.5% を「売り時」の乖離率とする．ただし，度数は対数表示である．

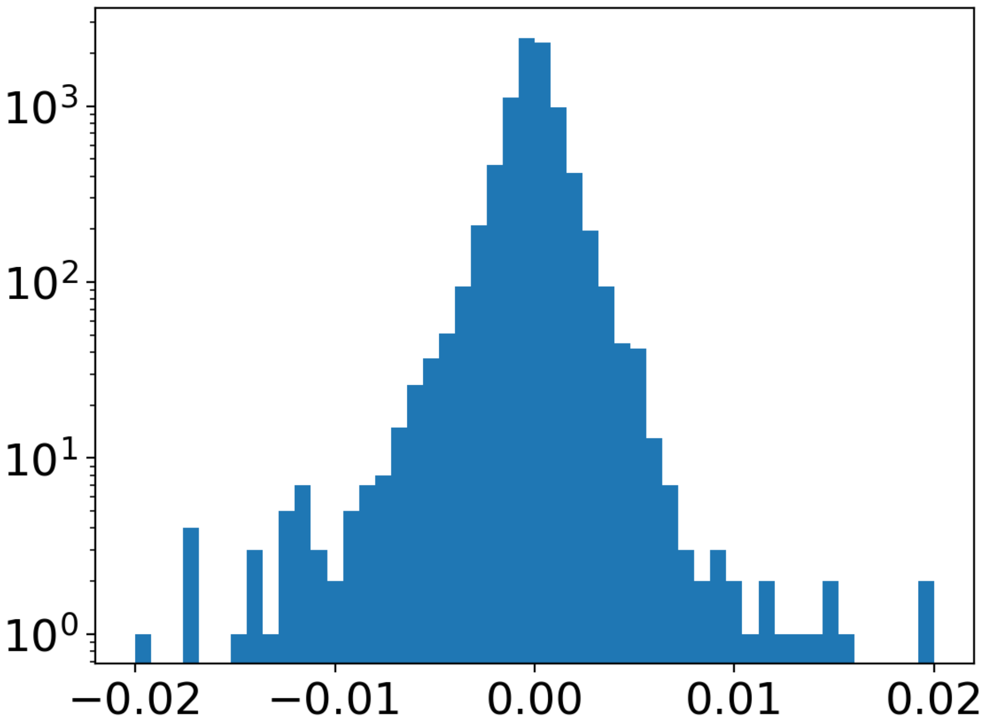


図 38　25 時点単純移動平均線乖離率分布

続いて，2018 年 10 月 01 日 00 時 00 分 00 秒から 2018 年 10 月 31 日 23 時 55 分 00 秒の 5 分ごとに観測された価格の，各時点における 100 時点単純移動平均線乖離率のヒストグラムを示す．25 時点単純移動平均線同様，以下に示す乖離率の分布の下位 2.5% の値を「買い時」の乖離率，上位 2.5% を「売り時」の乖離率とする．度数は対数表示である．

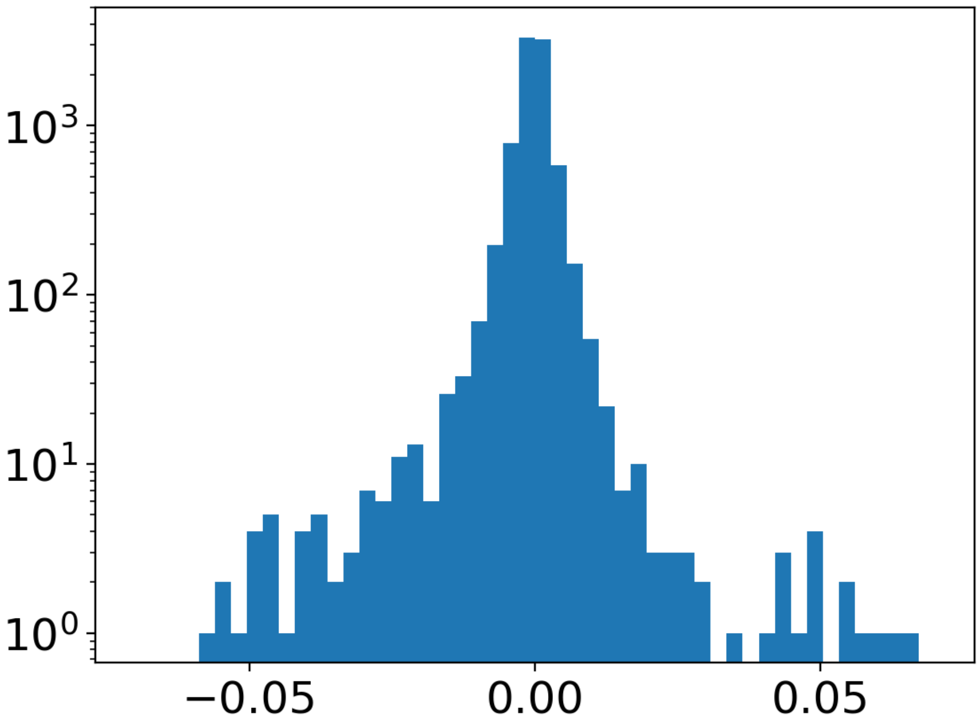


図 39　100 時点単純移動平均線乖離率分布

25 時点単純移動平均線乖離率の下位 2.5% の値 d_1，上位 2.5% の値 d_2 は，$d_1 = -0.0037, d_2 = 0.0033$ だった．また, 100 時点単純移動平均線乖離率の下位 2.5% の値 d_1, 上位 2.5% の値 d_2 は, $d_1 = -0.008, d_2 = 0.0063$ だった．したがって，25 時点単純移動平均線の場合は乖離率が-0.0037 となった時点を買いシグナル，0.0033 となった時点を売りシグナルとする．また，100 時点単純移動平均線の場合は乖離率が-0.008 となった時点を買いシグナル，0.0063 となった時点を売りシグナルとする．

2018 年 10 月 31 日 12 時 00 分 00 秒から 2018 年 11 月 30 日 23 時 55 分 00 秒における原系列データと 25 時点単純移動平均線，及び乖離率のチャートを以下に示す．はじめに全期間のものを示し，続いて期間を 2 分割したものを順に示す．

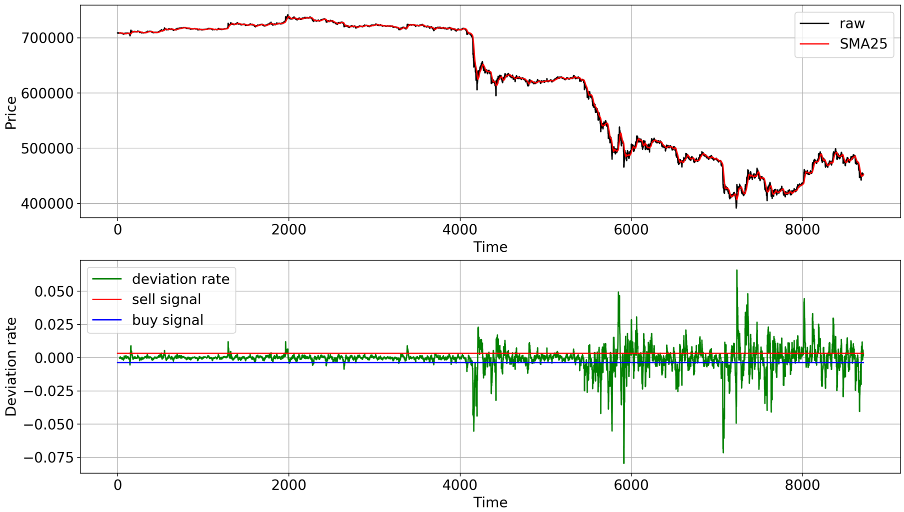


図 40　25 時点単純移動平均線乖離率チャート全期間

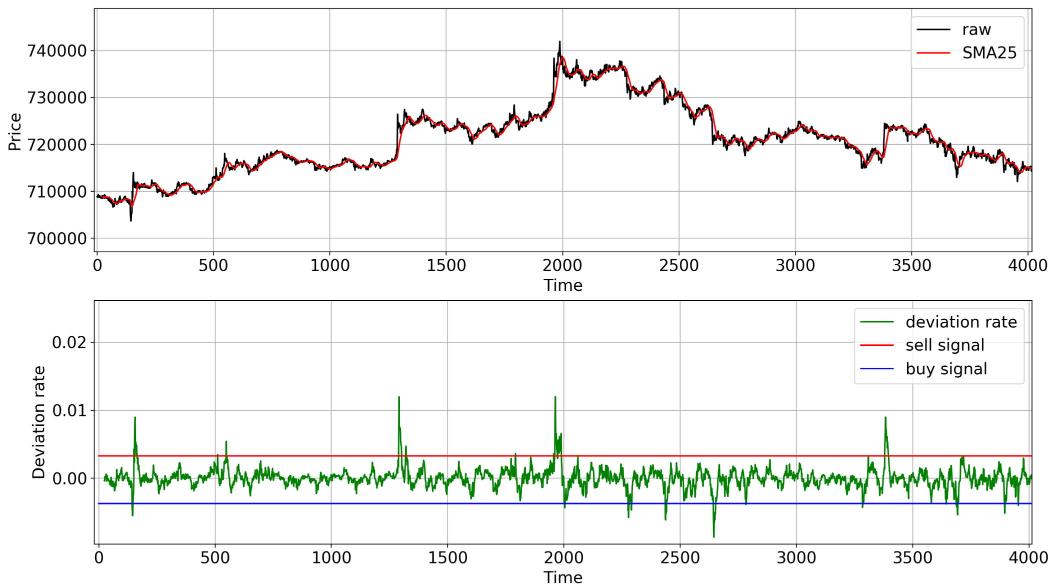


図 41　25 時点単純移動平均線乖離率チャート 1/2

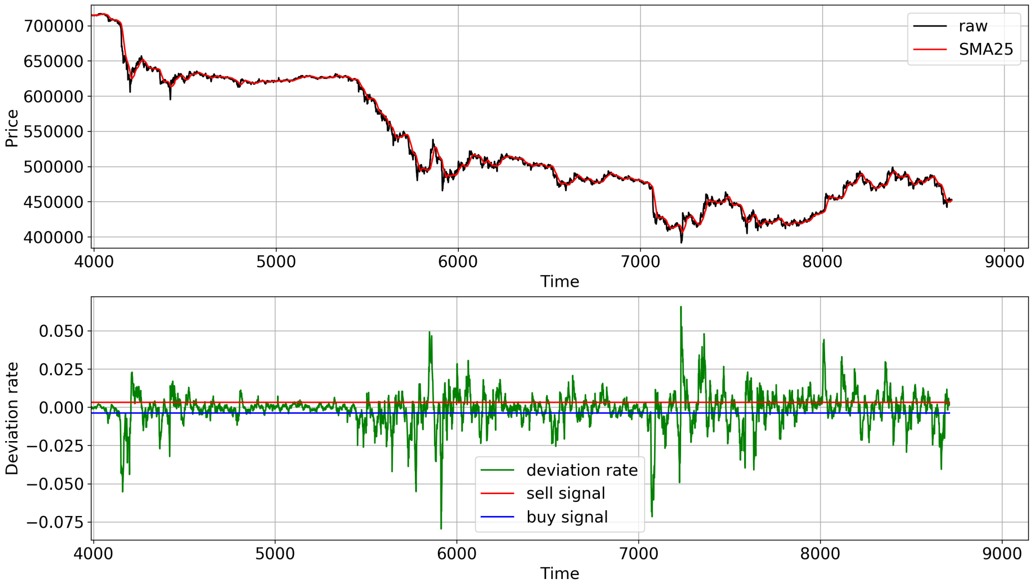


図 42　25 時点単純移動平均線乖離率チャート 2/2

25 時点単純移動平均線乖離率による売買の最終結果のログを以下に示す.

```
売買回数: 69
損益合計: -230973
合計収益率: 0.9942464168333077
勝利回数: 37
勝率: 0.5362318840579711
```

図 43　25 時点単純移動平均線乖離率売買結果

続いて，2018 年 10 月 31 日 12 時 00 分 00 秒から 2018 年 11 月 30 日 23 時 55 分 00 秒における原系列データと 100 時点単純移動平均線，及び乖離率のチャートを以下に示す．

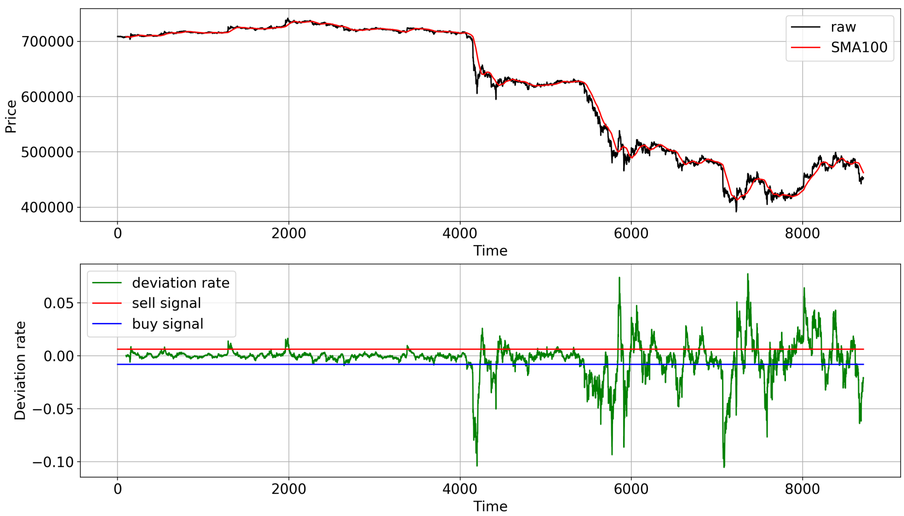


図 44　100 時点単純移動平均線乖離率チャート全期間

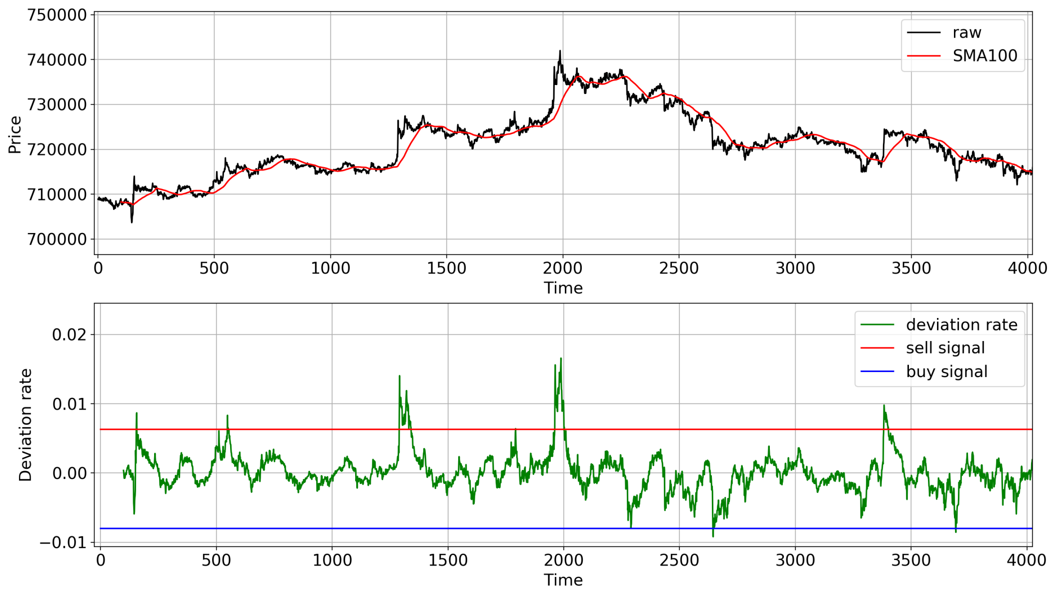


図 45　100 時点単純移動平均線乖離率チャート 1/2

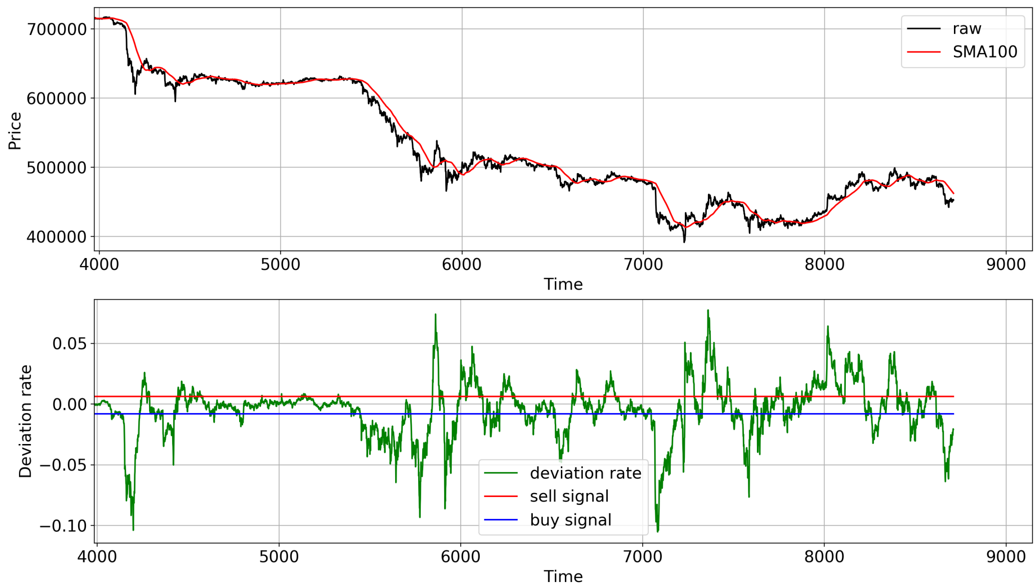


図 46　100 時点単純移動平均線乖離率チャート 2/2

100 時点単純移動平均線乖離率による売買の最終結果のログを以下に示す.

```
売買回数: 18
損益合計: -220376
合計収益率: 0.980431799310744
勝利回数: 9
勝率: 0.5
```

図 47　100 時点単純移動平均線乖離率売買結果

10　考察

10.1　単純移動平均線

5・25 時点単純移動平均線の売買結果は，図 23 の通りであり，最終的に利益をあげることはできなかった．しかし合計損失は約 2 万円であり，一回あたりの損失額は平均 85 円程度である．本研究ではロスカット等を行わなかったが，ロスカットのルールを定める等，売買する上で損失を最小限にした場合に，利益をあげることの可能性はあると言ってもよい．

具体的には，購入後に損失額が一定の水準になった場合は売りシグナルとなっていなくとも売却をするといったものである．売却基準の決定に関しては，例えば過去データの GC・DC 間の価格を一時点ごとに観測し，一度の売買に対する許容可能な損失を決定する方法が挙げられる．

また，25・75 時点単純移動平均線の売買結果は，図 30 の通りであり，最終的に 2894 円の利益をあげた．25・75 時点単純移動平均線の場合も．5・25 時点同様にロスカットのルールを定め，損失を最小限にすることが可能ならば，さらなる利益向上の余地がある．

以上より，5 分ごとに観測した BTC の価格に対して，単純移動平均線を利用し投資戦略を検討することは不可能であるとは言えない．買い・売りシグナルのみで決済をするだけでなく，GC・DC 間の価格及び価格変動を観測する等，再考の余地は十分にある．

10.2　MACD

期間 12・26，シグナル 9 の MACD による売買結果は，図 37 の通りであり，最終的に利益をあげることはできなかった．

MACD は，価格が横ばいで推移している明確なトレンドが観測されない相場 (ボックス相場) においては有用なテクニカル指標ではない．MACD ラインは，2 つの期間の異なる指数平滑移動平均の値の差分であるため，価格の変動が小さい相場が続いた場合は，MACD ラインの値も横ばいになる傾向にある．その場合は，MACD ラインの単純移動平均である MACD シグナルもいずれ横ばいになる．この状態では MACD ラインと MACD シグナルが非常に交差しやすくなるため，買い・売りシグナルである GC・DC も観測されやすくなる．さらに，MACD は移動平均線を利用したテクニカル指標であるために，実際の価格変動と比較して遅れて変動する傾向にあり，ボックス相場においては DC が観測された時点の価格 (売り時) が GC が観測された時点の価格 (買い時) を下回っている現象が発生しやすく，GC・DC にしたがった売買を行うと損失が発生しやすくなりやすい．

本研究ではいかなる場合においても GC・DC に従った売買を想定し損益を算出している．しかし上記の通りボックス相場において観測される MACD の GC・DC は有意なものとは到底言えず，GC・DC に従った売買は避けるべきである．従って，MACD を利用し投資戦略を検討する場合は MACD だけではなく，相場の方向性 (トレンド) を示すテクニカル指標も確認し売買を行う必要がある．

具体的には，例えばテクニカル指標の一つである DMI[17] を参照することが挙げられる．また，MACD ラインや任意の期間の移動平均線の傾き (現時点と前時点の値の差分) や変化率を観測するのも良い．

10.3 単純移動平均線乖離率

25 時点単純移動平均線乖離率を利用した売買結果は図 43，100 時点単純移動平均線乖離率を利用した売買結果は図 47 の通りであり，最終的に利益をあげることはできなかった．

しかし勝率は 25 時点・100 時点のいずれの場合も 50% 以上である．これは，一度の決済において発生した利益と比較して，損失が大きいことが考えられる．特に，今回検証した期間では，大きな下降トレンドに転換した時点で，大きな損失が発生している．ゆえに，このテクニカル指標を利用する場合にはロスカットのルールを定め損失を最小限にとどめる必要がある．

また，移動平均線乖離率を利用した売買は，トレンドが観測されている相場においては使用されるべきではない．なぜならば，図 42・46 のように，トレンドの転換が確認された場合には移動平均線乖離率の分散が急激に変化し，事前に定めた乖離率が買い・売りシグナルとして機能しなくなる現象が発生する傾向にあるためである．さらに，下降トレンドに転換した場合はロスカットをしなければ，損失が拡大しやすくなる．したがって，移動平均線乖離率を利用し投資戦略を検討する場合は，MACD 同様それ単体だけではなく，相場のトレンドを示すテクニカル指標も確認し，明確なトレンドが観測される場合は移動平均線乖離率を利用した売買は避けるべきである．

10.4 全体

本研究においては，実際に BTC の売買シミュレーションに採用した期間は 2018 年 10 月 31 日 12 時 00 分 00 秒から 2018 年 11 月 30 日 23 時 55 分 00 秒である．この期間の特徴として，一定期間ボックス相場であった後に，大きな下降相場となっていることが挙げられる．

まず，下降相場に転換する直前の 2018 年 11 月 14 日から過去 4 ヶ月間の BTC の 1 日ごとの価格 (終値) のチャートを示す．

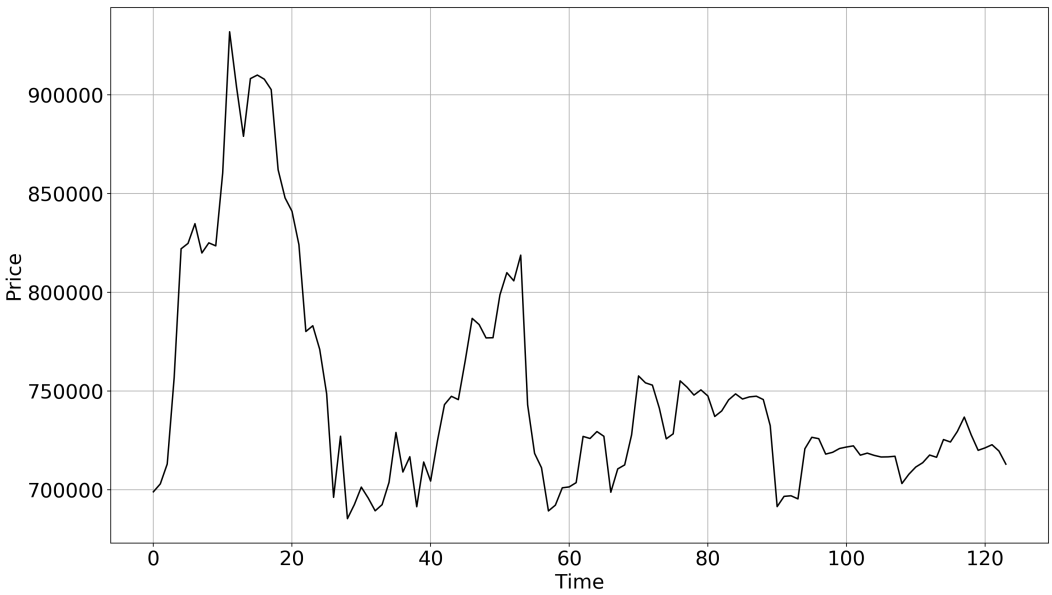


図 48　2018 年 7 月 14 日〜2018 年 11 月 14 日

図 48 より，該当期間の相場の形は先行き弱気の三角持ち合い [18] となっており，下降トレンドに転換する可能性が高い状態と言える．[18]

また，下降相場についてであるが，今回採用した期間中に，ビットコインキャッシュ (BCH) のハードフォーク [19] が発生するというニュースが事前に発表されていた．[20] このハードフォークは，クライアント開発チームの BitcoinABC と Nchain の立場の違いが加わった，いわゆる対立によるものであり，暗号通貨の価格下降リスクとなりうるものと言えた．

以上より，今回採用した期間はいずれ大きな下降トレンドが観測される可能性が高い状態だったと言える．

投資対象の商品を空売りした場合，空売り後に価格が下降した場合は利益をあげることができる．本研究では，空売りは行わないものとして検証を行ったが，上記のように価格の下降が予測される状況では，空売りを行うことによって価格の下降による損失を相殺する (リスクヘッジ) べきと言える．

11　今後の展望

以下の事柄が今後の展望の一例としてあげられる．

1. 他のテクニカル指標の導入・組み合わせによる投資戦略の検討
2. テクニカル指標パラメータ最適化
3. 原系列データ平滑化による対数差分の時系列解析

11.1　他のテクニカル指標の導入・組み合わせによる投資戦略の検討

テクニカル指標は今回採用したものの他にも多数存在する．例えばボックス相場における買われ過ぎ (今後価格が下降する可能性が高い)・売られ過ぎ (今後価格が上昇する可能性が高い) 度合いを測り買い・売りシグナルを示す指標として，RSI やストキャスティクスが存在する．

また，MACD 及び単純移動平均線乖離率の考察で述べた通り，テクニカル指標を複数組み合わせて売買を行うと，損失を小さくすることが可能な場合が存在する．

他のテクニカル指標を導入し，複数のテクニカル指標を参照し，可能な限り損失を小さくするような投資戦略を検討する余地は十分にあると言える．

11.2　テクニカル指標パラメータ最適化

多くのテクニカル指標には，それぞれ可変のパラメータが存在する．そのテクニカル指標に対してのパラメータを調整し，利益を最大化するといった研究の余地があると言える．

一例として，移動平均線の可変のパラメータである係数・期間の調整による利益の向上の可能性について記載する．

任意の時点 t における移動平均は，一般的には時系列データ $X_1, X_2, ...$ の任意の期間 p 中の値に何らかの係数 $\phi_1, \phi_2, ..., \phi_p$ をかけたものの総和である．すなわち，時系列データ $X_1, X_2, ...$ の t 時点における期間 p の移動平均 $MA(t,p)$ は，

$$MA(t,p) = \sum_{i=0}^{p-1} \phi_{i+1} X_{t-i} \tag{20}$$

と表現される．

係数が一定値ではない移動平均の一種に，線形加重移動平均といったものが存在する．t 時点における期間 p の線形加重移動平均 $WMA(t,p)$ は，

$$\begin{aligned} WMA(t,p) &= \frac{pX_t + (p-1)X_{t-1} + \cdots + X_{t-p+1}}{p + (p-1) + \cdots + 1} \\ WMA(t,p) &= \frac{\sum_{i=0}^{p-1}(p-i)X_{t-i}}{\frac{1}{2}p(p+1)} \end{aligned} \tag{21}$$

と算出される．

以下に，2018 年 11 月 27 日 20 時 00 分 00 秒から 2018 年 11 月 29 日 18 時 00 分 00 秒の BTC 価格に対する期間 $p = 25$ の単純移動平均線及び線形加重移動平均線を描画したものを示す．

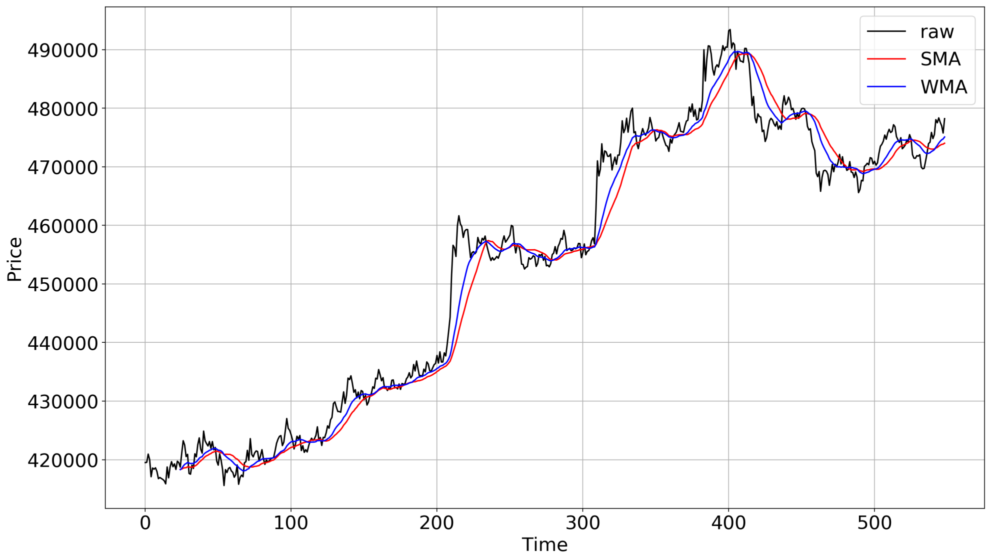


図 49　2 種類の移動平均線比較

凡例の raw は原系列データ，SMA は単純移動平均線，WMA は線形加重移動平均線である．図 49 より，線形加重移動平均は単純移動平均と比較し，より原系列データに近い値となっていることがわかる．線形加重移動平均による GC・DC は，単純移動平均と比較してより早くトレンドの転換に対応することが可能であり，より利益を増加・損失を減少させることが期待される．

一方で，移動平均の計算に使用する値のうち，最近の値にかける係数を過剰に重くした場合は，価格の変動にも過剰に反応し，有意とは言えない GC・DC も観測されやすくなる．

したがって，移動平均の係数や期間を調整し最適化することが可能ならば，GC・DC をはじめとした移動平均線のテクニカル指標を利用した投資戦略を検討する上で利益の最大化が期待される．

11.3　原系列データ平滑化による対数差分の時系列解析

本研究においては，原系列データの対数差分に対して統計的性質の調査を行ったところ，有意な自己相関が観測されず，確率分布の推定を行うことは不可能だった．しかし，原系列データの 2 時点単純移動平均線の対数差分に対して自己相関を調査すると，有意な自己相関係数が観測される．

本研究では時間の都合上，それに対しての時系列解析及び確率分布の推定をするに至らなかった．しかし原系列データの 2 時点単純移動平均線をはじめとした，原系列データを平滑化したものの対数差分を時系列解析し，確率分布の性質を推定する研究の余地はあると言ってもよい．ただし過剰な平滑化を行うと，原系列データとは全く別の性質のデータになる可能性が高く，その場合は解析した結果は無意味なものとなる．

12 謝辞

本論文は筆者がサレジオ工業高等専門学校情報工学科 5 年次に在籍中の研究成果をまとめたものである．准教授佐藤豊先生には指導教官として本研究の実施の機会を与えて戴き，その遂行にあたって終始，ご指導を戴いた．ここに深謝の意を表する．

13 付録

今回使用したプログラム及びデータは，GitHub 上に公開している．[21]

参考文献

[1] 日本のビットコインが使えるお店 (ビットコイン決済対応店舗) — Bitcoin 日本語情報サイト https://jpbitcoin.com/shops

[2] 矢部大輔・木村昌臣 (2008) 「データマイニングによる株価予測システムの開発」,『全国大会講演論文集 第 70 回 (ソフトウェア科学・工学)』pp.281-282.

[3] 砂田吉一・橋本次郎 (1989) 「多変量解析による為替レートの要因分析 (第 2 回計算機統計学会シンポジウム報告 経済と経営の実証分析 : 計算機の利用)」,『計算機統計学 2(1)』 pp.97-98, 日本計算機統計学会

[4] Market Data REST API https://cryptowatch.jp/docs/api

[5] MySQL https://downloads.mysql.com/archives/installer/

[6] The Comprehensive R Archive Network https://cran.ism.ac.jp

[7] mysqlclient・PyPl https://pypi.org/project/mysqlclient/

[8] mrjbq7/ta-lib: Python wrapper for TA-Lib (http://ta-lib.org/). https://github.com/mrjbq7/ta-lib

[9] R で計量時系列分析：はじめに覚えておきたいこと - 六本木で働くデータサイエンティストのブログ https://tjo.hatenablog.com/entry/2013/07/04/190139

[10] 自己共分散と自己相関 — 時系列データにおけるある時刻の観測値とその過去のある時刻の観測値との相関 https://stats.biopapyrus.jp/time-series/acf.html

[11] 横内大介・青木義充 (2018).『現場ですぐ使える時系列データ分析』. 技術評論社, 223p

[12] 馬場真哉 (2018).『時系列分析と状態空間モデルの基礎 R と stan で学ぶ理論と実装』. プレアデス出版, 350p

[13] AIC による AR モデルの次数決定法 [ストレスと自律神経の科学] http://hclab.sakura.ne.jp/stress_nervous_ar_aic.html

[14]【統計学】ホワイトノイズとは？ https://tkstock.site/2017/04/05/2017-04-05-000000/

[15] MACD の見方と使い方は？億トレーダーが初心者向けに解説！ https://fx-megabank.com/technical-analysis/macd/

[16] 空売り｜初めてでもわかりやすい用語集｜SMBC日興証券 https://www.smbcnikko.co.jp/terms/japan/ka/J0320.html

[17] DMI の見方・使い方 — テクニカル分析指標 — 指標の見方・使い方 — 投資のノウハウ — 株の達人 https://www.sevendata.co.jp/shihyou/technical/dmi.html

[18] 三角持ち合い｜初めてでもわかりやすい用語集｜SMBC日興証券 https://www.smbcnikko.co.jp/terms/japan/sa/J0479.html

[19] ハードフォークとは - とってもやさしいビットコイン https://www.tottemoyasashiibitcoin.net/entry/2016/10/14/113708

[20] 11 月 15 日の BCH ハードフォークに海外 2 大取引所 Coinex と Bitasiaex が対応方針を発表 https://coinpost.jp/?p=44320

[21] technical-college-graduation-research-mirror https://github.com/yyk458/technical-college-graduation-research-mirror