

前回同様、[ウェブサイト](#)から入手した日経平均株価のデータを分析する。2018/1/4 から 2021/6/18 までの 841 件のデータを取得した。

前は株価そのものの予測を行うことを目標としていた。しかし効率的市場仮説の考えに基づけば、株価は将来に対するあらゆる情報を織り込み済みなので株価自体の予測は現実的に難しいと考えた。

そのため今回は株価の下落とボラティリティの関係を調べてみようと思う。

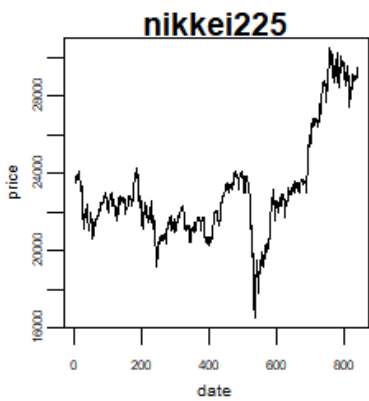


図 1 元データ

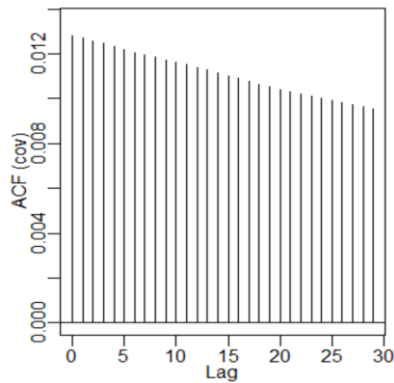


図 2 自己共分散

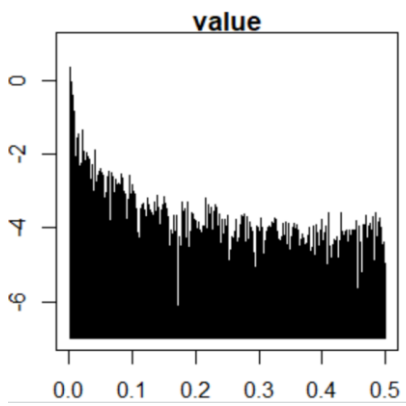


図 3 ピリオドグラム

日経平均データは分散が時間変化する時系列データであると考えられる。そのため非ガウス型平滑化を用いて時変分散を推定する。

平滑化に用いるシステムノイズと観測ノイズの確率分布を、表 1 の①～③の 3 通り仮定する。

表 1 平滑化のノイズ分布

	システムノイズ	観測ノイズ
①	正規分布	2 重指数分布 (非ガウス)
②	正規分布	正規分布 (ガウス型)
③	コーシー分布	2 重指数分布 (非ガウス)

それぞれの場合の対数尤度及び AIC は表 2 のようになった。

表 2 ①~③の AIC

	対数尤度	k	AIC
①	-856.641	2	1717.28
②	-874.595	2	1753.19
③	-795.604	4	1599.21

以上から AIC 最小は③のモデル、つまりシステムノイズにコーシー分布、観測ノイズに 2 重指数分布を仮定したモデルであることがわかる。よってモデル③を用いて時変分散を推定する。

なお、①~③の $\log \sigma^2$ の事後分布は図 4 のようになった。AIC が大きくなるほど起伏の少ないグラフとなっていることが見て取れる。

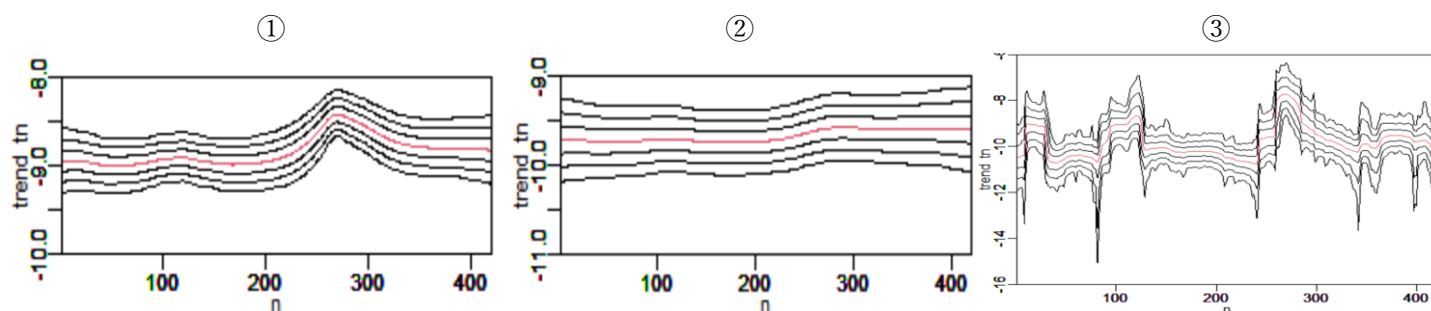


図 4 表 1 の①②③におけるそれぞれの時変分散の推定

モデル③を用いて株価データのボラティリティの推定を行うと図 5 のようになった。ボラティリティの大きい部分と株価の大幅な下落時点が重なっている。元データの 50,250 時点あたりの小さな下落でも時変分散の推定値は大きくなっている。よって日経平均株価は株価の上昇・レンジ期間より下落期間のほうがボラティリティは大きくなる傾向にあることがわかった。

モデル③の事後分布の 3 次元プロットは図 6 のようになった。非ガウス型平滑化なので時刻に対してトレンドが不連続的に変化している。

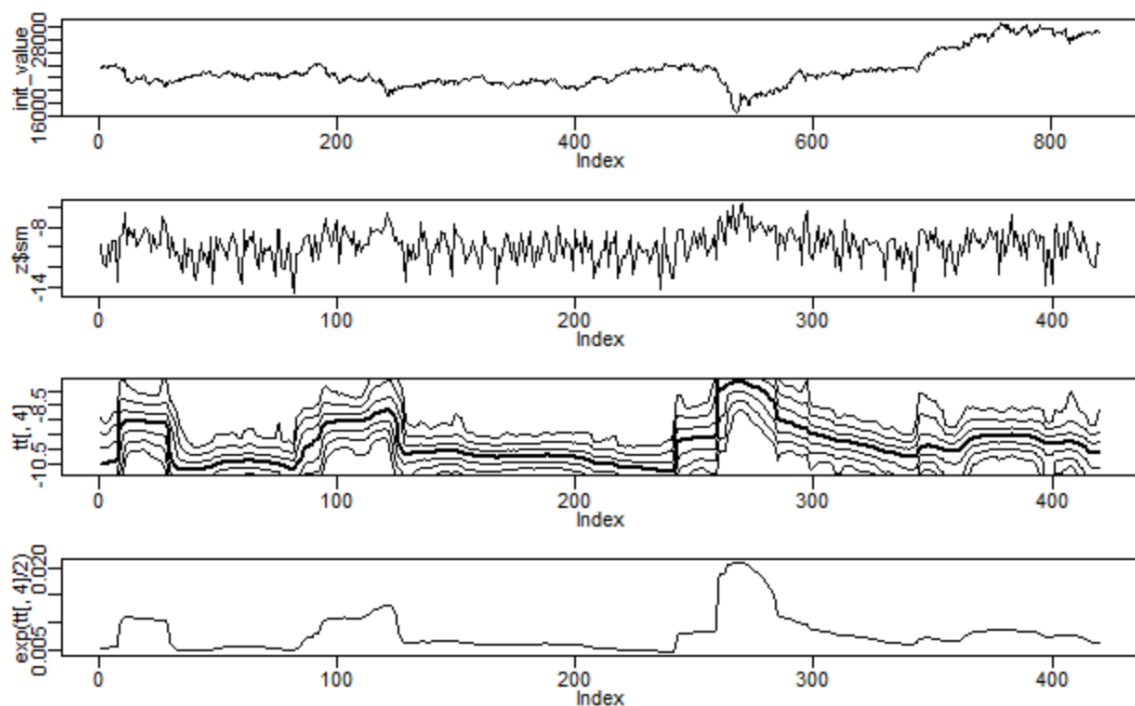


図5 モデル③を用いた株価のボラティリティ推定、上から元データ(対数差分変換なし)、変換後のデータ、 $\log \sigma^2$ の事後分布、ボラティリティの推定値。

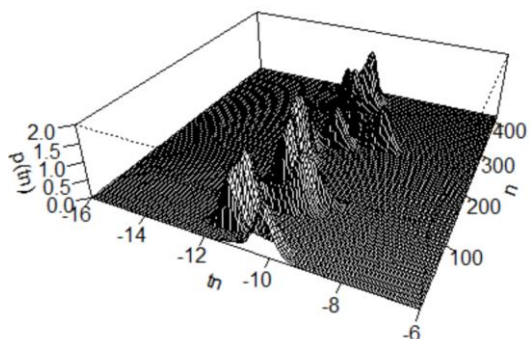


図6 モデル③の事後分布

● 総合的な考察

1. 設定した課題はそもそもデータに基づく分析ができるようなものであったか。また、分析目的にあった適切なデータを取得できたか。

株価の下落とボラティリティの関連性は、データに基づく分析が可能であった。日経平均株価のデータは Web 上で csv 形式としてダウンロード可能であったので、適切なデータを容易に取得できた。

2. 解析の結果，期待したような結果や想定外の面白い知見が得られたか.

期待通り、株価は下落時に大きなボラティリティとなることが分かった。**AIC** 最小となったのはモデル③（システムノイズ：コーシー分布、観測ノイズ：2重指数分布）のときであった。株価はコロナショックなどで簡単に大幅下落することを考えると、正規分布より裾の重いコーシー分布モデルが採択されるのは自然な結果だと思った。