

ベイジアン VAR を適用したマクロ経済情報 の波及分析による伝播予測

蒲田 涼馬 (Ryoma Gamada)
u020010@st.pu-toyama.ac.jp

富山県立大学情報システム工学科 4 年

December 26, 2023

背景

2/9

背景

近年、時系列データや金融指標を用いたテクニカル分析やファンダメンタル分析、ニュース情報を用いたテキスト分析など様々な形で株価の予測が行われている。
株価の動向を予測することは投資の意思決定への支援、リスク管理などで重要になる。

目的

業種間分析を行うことで経済の動向を把握し、投資判断をサポートする。
また、市場変化への戦略策定も支援する。

従来研究

3/9

ベイジアンネットワークを用いた本邦市場の伝播構造の分析

本邦の株式市場におけるストレス事象発生時における業種間の伝播構造をベイジアンネットワークを用いて把握した研究.

BVAR モデルを日本経済のマクロデータに適用した場合の実証結果を示した研究. (5 つの時系列データ使用)

これらの研究ではパッケージを用いて BVAR 解析, 考察を行っているが有用性などの判定はしていなかった.

システムの流れ

4/9

システムの流れ

主にサンプリング, LPA(潜在プロファイル分析), 何らかの手法によるモデルの作成, RF の適用, 可視化の 5 ステップからなる予定.

モデル作成

サンプリングしたデータの欠損値を導出するために目的関数のモデルを作成する必要がある.

ファジィニューロと回帰を試し, 良い結果の方を使う.

データ収集

データ収集に用いるサイト

株価データを分析に用いるために、各業種間の株価時系列データを取得する必要がある。

今のところ見つかっている無償で株価のデータを収集できるサイトは「みんかぶ」。

ここでは国内上場企業(東証33業種)の株価を日足時系列データ(土日祝を除く)で取得することができる。

スクレイピングについて

「みんかぶ」はcsv形式でのデータは公開していないのでスクレイピングによってデータを収集する必要がある。

また、必要なだけの時系列データを取得するためにはクリックする必要があるため、スクレイピングでクリックを使うことができるPythonのライブラリ「selenium」を用いることにした。

可視化について

6/9

可視化

本研究では投資の意思決定支援、また社会動向の理解を支援するために可視化を行う。

可視化については以前 BVAR の際に用いていた 3D グラフによる可視を使う。

システムによって出力された数値を csv に格納し、それから有向グラフを作成する。

またノードをクリックした際には最新のデータおよび過去のデータを確認できるようにする。



変数とデータについて

7/9

変数

変数はマクロ経済データ(約120個)を扱う。

ミクロではなくマクロの経済データを扱うのはマクロの方がデータを持ってきやすいから。

スクレイピング

日本銀行時系列サイトから自動でデータを持ってくるシステムを作成した。

データは csv でダウンロードしたもの、またページによってはさらに高度なスクレイピングが必要になってくる。

csv でダウンロードできるものは1つの csv ファイルにまとめて扱う予定。

実装

同定式導出の実装

適当に作成した特徴量 $x1 \sim x5$ から目的変数 $y1 \sim y5$ の同定式の導出を行うシステムを実装した。 $y1, y2$ の同定式は掛け算と足し算のみをした単純なものであるため、正しい同定式は導出できたが、指數を扱った $y3 \sim y5$ についてはまったく良い結果が得られなかった。

正解データの例:

$$y1 = x1 + x2 + x3 + x4 + x5$$

$$y3 = x1 + x2 + x3^2 + x4 - 400$$

$x1$	$x2$	$x3$	$x4$	$x5$	$y1$	$y2$	$y3$	$y4$	$y5$
12	8	21	22	7	70	182	96	662	1010
13	6	22	23	6	70	183	142	704	1099
11	9	25	23	8	76	198	284	735	1293
15	5	24	25	5	74	196	240	809	1319
14	6	23	21	5	69	190	187	623	1069
17	8	26	21	6	78	225	340	654	1297
14	9	22	20	8	73	201	140	591	1005
14	9	26	24	7	80	215	340	791	1426
14	7	25	23	6	75	203	287	728	1293
13	8	25	22	9	77	207	285	687	1256

図 2: デモデータ

$$\begin{aligned}
 y1 &= 4.7263956413504801e-14 + 1.000 * x1 + 1.000 * x2 + 1.000 * x3 + 1.000 * x4 + 1.000 * x5 \\
 y2 &= 0.0 + 5.000 * x1 + 3.000 * x2 + 4.000 * x3 + 0.000 * x4 + 2.000 * x5 \\
 y3 &= 972.0012155743054 + 1.660 * x1 + 2.070 * x2 + 47.408 * x3 + 1.670 * x4 + -0.678 * x5 \\
 y4 &= -524.0409016533799 + 2.453 * x1 + 5.308 * x2 + 5.098 * x3 + 45.098 * x4 + 0.529 * x5 \\
 y5 &= -1496.059823205808 + 10.114 * x1 + 11.378 * x2 + 67.583 * x3 + 48.368 * x4 + 2.858 * x5
 \end{aligned}$$

図 3: 結果

まとめ

9/9

進捗まとめ

為替データではなく、株価の時系列データについての研究をすることになった

データ収集システムの実装が進んだ

回帰による複数の目的関数に対しての同定式の導出を実装してみた

今後やること

モデル作成システムを完成させる

可視化システムを作成 (flask でサーバー上に実装)

数値実験と本論への取り組み