

中国の省レベル経済ネットワークの研究

南川 高範

12月19日（土）

これまでの研究

これまでの視点

- 物価
- 景気後退
- ホットマネー

これまでの研究

現在の視点

- 地域経済
- 国内経済の潜在的ネットワークの可視化
→ 今回紹介する研究（感染症の感染経路を推定し、そこから経済ネットワークについて考察する研究）
- 潜在的ネットワークの活用

研究の前提

ウイルスの流行

- 2020年1月6日日本の厚生労働省が報道機関向けに原因不明の感染性肺炎が発生していることを発表
- 感染症の名前はCOVID-19、その原因としてSARS-COV-2と称される新型コロナウイルスが特定

研究の前提

ウイルスの流行

- 日本では1月15日に第1例目の感染者が確認され、現在も世界中に流行

研究の前提

感染症と経済

- 感染予防措置として各国から日本への移動を禁止する国は2020年9月時点で114か国
→ 日本政府も国民の出国に対して中止勧告を促す措置
- 感染リスクがもたらす物理的、心理的な作用により経済活動が委縮

研究の前提

感染症と経済

→ マスクの不足などにみられる資源配分のゆがみ

- 日本の実質GDP成長率は2020年4月から6月にかけて -8.3% という水準を示し、その後7月から9月にかけて 5.3% に上昇

研究の前提

感染症と経済

→ 感染症により生産活動が大きく攪乱されている

研究の前提

感染の原因

COVID-19は現時点では、人間同士の間でのウイルスの移動により拡大しているとみられている

→ 感染者数の多い都市では感染が広がりやすく、感染者数の多い都市からの人口移動が多い都市でもまた、感染者が増加すると考えられる

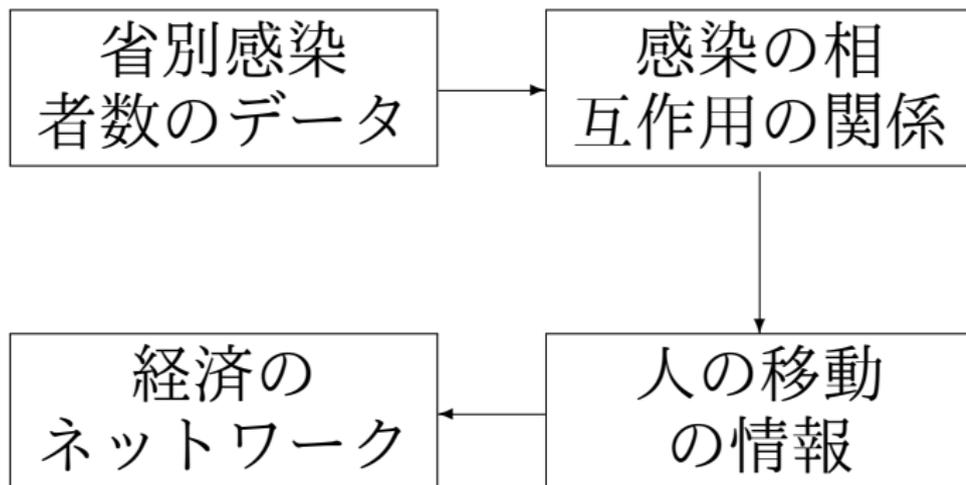
研究の前提

研究の目的

中国における省レベルの経済ネットワークの様子を明らかにすること

→ そのために中国国内における Covid-19 感染者数のデータを用いる

研究の前提



データとモデル

仮定に基づく感染のモデル

$$\Delta I_{it} = \sum_{i=1}^{31} w_{ij} I_{j,t-p} + \text{error}_{it}$$

- I_{it} : i 省における t 時点の累積感染者数
→ Δ は差分を意味する
- p はラグ次数

データとモデル

仮定に基づく感染のモデル

$$\Delta I_{it} = \sum_{i=1}^{31} w_{ij} I_{j,t-p} + \text{error}_{it}$$

- w_{ij} は i 省と j 省の間の結びつきの強さを表すウェイト
- error_{it} は誤差項

データとモデル

感染モデルの行列表記

$$\Delta I_t = W I_{t-p} + \text{error}$$

$$I_t = \begin{bmatrix} I_{\text{北京},t} \\ I_{\text{天津},t} \\ \vdots \\ I_{\text{西藏},t} \end{bmatrix}, \quad W = \begin{bmatrix} W_{\text{北京}, \text{北京}} & W_{\text{北京}, \text{天津}} & \cdots & W_{\text{北京}, \text{西藏}} \\ W_{\text{天津}, \text{北京}} & W_{\text{天津}, \text{天津}} & \cdots & W_{\text{天津}, \text{西藏}} \\ \vdots & \vdots & \ddots & \vdots \\ W_{\text{西藏}, \text{北京}} & W_{\text{西藏}, \text{天津}} & \cdots & W_{\text{西藏}, \text{西藏}} \end{bmatrix}$$

データとモデル

感染モデルの行列表記

$$\Delta I_t = W I_{t-p} + \text{error}$$

- 実際には W の対角要素を 0 とし、空間ウェイト行列とみなす
- W は人の移動の活発さを表す指標となり、経済的な結びつきの強さと解釈

データとモデル

データの制約

- 使用するデータは日本経済新聞がウェブ上で公開している各省感染者数のデータ
→31の省の累積感染者数を2020年1月22日から4月3日まで公開
- 2020年1月23日から4月8日までは湖北省武漢において都市封鎖を実施

データとモデル

データの制約

→ 武漢市を出入りするような人の移動が遮断

- 当該データが対象とする期間は武漢の出入りがない状態

データとモデル

武漢の感染者数を除外

- 中国国家衛生健康管理委員会が湖北省の感染者数をウェブサイト上で武漢市とそれ以外で分けて公開
→ 公開期間は2月13日以降

データとモデル

武漢の感染者数を除外

- 1月22日時点で中国全体の感染者数は440人だったのに対して、3月31日時点では同数字が81518人に増加
→3月31日時点で武漢の感染者数が50007人

データとモデル

武漢の感染者数を除外

- 本研究の対象はこの武漢の感染者以外の約3万人
→ 武漢の都市封鎖から2週間以上がたった2月13日から3月31日までの期間が対象

データとモデル

モデルの推定方法

- 新規感染者数の決定モデル

$$\Delta I_t = W I_{t-p} + \text{error}$$

を推定する際に問題となるのがサンプルサイズと係数行列の次元の問題

→ 係数パラメータの数は $31 \times 30 = 930$

データとモデル

モデルの推定方法

- 空間ウェイト行列は要素に0を多くもつことが予想され、小標本高次元データのためのモデル推定手段の一つである

Least Absolute Shrinkage Selection

Operator (LASSO) 型の推定を実行

→ 変数選択をしながら残差二乗和を最小にしていくという推定方法

データとモデル

モデルの推定方法

- 930の変数に対して0に近い係数を与えるものを除外し、選択された変数のパラメータを操作して目的関数の最適化

$$L(\mathbf{W}) = (\Delta \mathbf{I}_t - \mathbf{W} \mathbf{I}_{t-p})' (\Delta \mathbf{I}_t - \mathbf{W} \mathbf{I}_{t-p}) - \lambda \|\mathbf{w}_i\|_1$$

データとモデル

モデルの推定方法

損失関数 L を最小化する W を特定、但し、 w_i は行列 W の要素

検証結果

ラグ次数の決定

- ここで使用する感染者数のデータは日足データであり、累積感染者と新規感染者の間に数日のラグを伴うと考えられる
→ 厚生労働省の発表によるとウイルスの潜伏期間は1から12.5日（多くは5~6日）とされている

検証結果

ラグ次数の決定

- 実際には発症から検査、確認、発表までに数日を要することも考えられることから、 $p = 1$ から $p = 20$ というラグを想定し検証

→ ラグ次数 13 日以降で、非ゼロ要素が急増し、12 日以内のラグでは他地域の感染の効果が十分に反映されていない

検証結果

ラグ次数の決定

- ここでは、最も非ゼロ要素が多かった14日のラグ次数を検証の対象としてラグ次数を選択

ラグ次数 (p)	1	2	3	4	5	6	7	8
非ゼロ要素の数	20	30	22	22	10	12	12	6
ラグ次数 (p)	9	10	11	12	13	14	15	16
非ゼロ要素の数	10	12	8	8	74	84	82	78
ラグ次数 (p)	17	18	19	20				
非ゼロ要素の数	68	62	62	60				

検証結果

空間ウェイト行列の特定

- LASSO型推定の一つである Elastic Net 推定により、推定されたパラメータ行列 \hat{W} を空間ウェイト行列として解釈するために対称行列に変換

$$\hat{W}_s = \frac{1}{2} \hat{W} \cdot \hat{W}'$$

検証結果

空間ウェイト行列の特定

- 31の省について、非ゼロ要素の中央値は2であり、非ゼロ要素が4つ以上ある省の数は10
→非ゼロ要素が4以上ある省を経済ネットワークの結節点と解釈

検証結果

空間ウェイト行列の特定

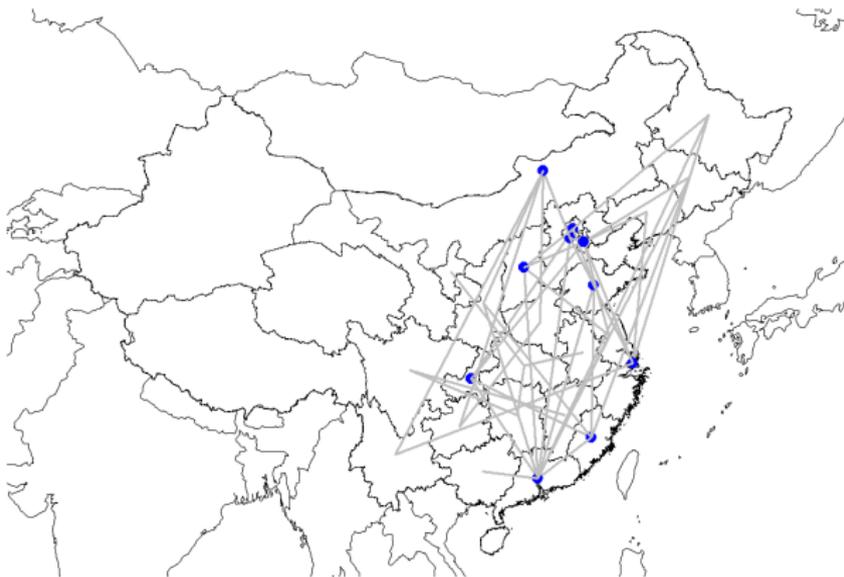
- 北京、天津、上海、福建、山東、広東、重慶という一人当たり所得の高い地域が結節点となっている傾向がある一方、所得の高い浙江省では非ゼロ要素が0となり、所得だけで説明できるわけではない

検証結果

空間ウェイト行列の特定

- 浙江省の非ゼロ要素がゼロである理由として、人口当たりの感染者数が多いことが考えられる
- 河北省は、一人当たり所得は高くないものの、人口が上位10位に入る省

検証結果



検証結果

結果の解釈

- 非ゼロ要素で関係が示されている省間を線で表示
- 結節点の中でも5つ以上の省と線で結ばれている省は、天津、内モンゴ、上海、福建、広東の5つ

検証結果

結果の解釈

- 省界を接する省同士が直接ネットワークを形成するのではなく、結節点を介してネットワークで結ばれるという特徴
→ 特に東北地方として分類される内蒙古、黒龍江、吉林、遼寧は直接線で結ばれていない

検証結果

結果の解釈

- ここでの結果は、COVID-19の感染経路を基に算出した、人の移動の多寡を表すもの
→ 平時における移動と異なる、湖北省については、武漢を除いた人の動きを基にしている

検証結果

結果の解釈

- 感染リスクを考慮しても、移動しなければならぬほどの強い経済的な結びつきや、商業上の慣習に基づく人の移動を反映したものであると解釈できる
→ 距離や物理的な接地以外の要因による、財・人の移動を説明

検証結果

結果の解釈

- 実質GDPや物価、貨幣供給などの経済変数について、空間的自己相関であるとか、空間計量経済モデルを推定する際に、一般的には距離や省界の共有などの既知の情報を活用

検証結果

結果の解釈

→ グラビティモデルに代表されるような、2地点間の財や人の移動については経済的な結びつきの強さの情報の方が実態を反映

Otto, Phillip and Rick Steinert. (2018) "Estimation of the Spatial Weighting Matrix for Spatiotemporal Data under the Presence of Structural Breaks", unpublished paper (URL)

<https://arxiv.org/abs/1810.06940>

安道知寛 (2014) 「高次元データ分析の方法-Rによる統計的モデリングとモデル統合」『朝倉出版』
中国国家统计局季度数据 (実質 GDP 成長率)

https:

[//data.stats.gov.cn/easyquery.htm?cn=B01](https://data.stats.gov.cn/easyquery.htm?cn=B01)

中国国家统计局分省年度数据 (各省人口、付加価値生産額データ) <https://data.stats.gov.cn/easyquery.htm?cn=E0103>

GADM（中国シェープファイル）

https://gadm.org/download_country_v3.html

Natural Earth（世界地図シェープファイル）<https://www.naturalearthdata.com/downloads/>

厚生労働省「中華人民共和国湖北省武漢市における原因不明肺炎の発生について」（報道機関向け発表）

https://www.mhlw.go.jp/stf/newpage_08767.html

厚生労働省「新型コロナウイルスに関連した肺炎の患者の発生について（1例目）」（報道機関向け発表）

https://www.mhlw.go.jp/stf/newpage_08906.html

https://www.mhlw.go.jp/stf/newpage_08906.html

https://www.mhlw.go.jp/stf/newpage_08906.html

厚生労働省広報「新型コロナウイルスを防ぐには」

<https://www.mhlw.go.jp/content/10900000/000596861.pdf>

日本経済新聞「新型コロナウイルス世界マップ」

<https://vdata.nikkei.com/newsgraphics/coronavirus-world-map/>

日本経済新聞「新型コロナウイルス感染中国マップ」

<https://vdata.nikkei.com/newsgraphics/coronavirus-china-map/>

日本経済新聞「GDP 実質 27.8 % 減、4~6 月年率 戦後最大の下げ」<https://www.nikkei.com/article/DGXMZ062699240X10C20A8MM0000/>

日本経済新聞「新型コロナウイルス感染中国マップ」

<https://vdata.nikkei.com/newsgraphics/coronavirus-china-map/>

BBC News「中国・武漢市、都市封鎖を解除2カ月半ぶり市民ら移動」

<https://www.bbc.com/japanese/52210389>

中国国家衛生健康管理委員会「発生通知」http://www.nhc.gov.cn/xcs/yqtb/list_gzbd_10.shtml?