

2021年2月26日 全7頁

データサイエンスとマクロ経済分析①

新局面を迎えるナウキャストイング

新型コロナが促すマクロ経済分析へのデータサイエンスの本格的導入

経済調査部 溝端 幹雄

[要約]

- 本シリーズでは、近年のマクロ経済分析・予測におけるデータサイエンス（ビッグデータ・オルタナティブデータのような新しいデータや、機械学習のような新しい分析手法）の活用状況を概観することにしたい。
- 新型コロナウイルス感染症による影響が長期化する中、急激な経済環境の変化を迅速に把握に対応するためナウキャストイングモデルに耳目が集まっている。ただし、今回のコロナ禍で話題になっているのは、現在のナウキャストイングモデルが足元の急激な変化を十分に捉えきれなかったことだ。その原因として、公表までにタイムラグのあるマクロ経済データを使っていることや、モデルの構造上、急激な経済悪化（非線形的な動き）をモデルで追認できないことが挙げられている。
- こうした課題に対して、ナウキャストイングモデルに電力消費データなどのリアルタイムで取得できる高頻度データを用いて、高精度かつリアルタイムでマクロ経済動向を捉える研究が出てきている。さらに、これまでビッグデータを用いたマクロ経済予測のパフォーマンスが期待されたほど良くなかったこと、モデルの構造がブラックボックスなために予測の運用上で欠点があったことなどの課題も、次第にその解決策が提示されつつある。
- もちろん、データサイエンスをマクロ経済分析に適用するには超えるべき課題も多い。しかしながら、海外の様々な研究が示すように、ビッグデータや機械学習などのAI的手法がマクロ経済分析においても大きな期待を寄せられている点は変わらない。今後、マクロ経済分析でも地に足の着いた形でデータサイエンスの活用が本格化するものと期待される。

本シリーズでは、近年のマクロ経済分析・予測におけるデータサイエンス¹（ビッグデータ・オルタナティブデータのような新しいデータや、機械学習などのような新しい分析手法）の活用状況を概観することにした。

1. コロナ禍で注目されるナウキャストイング

足元の経済状況の早期把握が目的

新型コロナウイルス感染症による影響が各方面で長期化している。昨年春には各国でロックダウンが実行に移されたこと、そして日本でも緊急事態宣言が発出されたことで、多くの経済・社会活動が停止に追い込まれた。それにより国内外のマクロ経済環境も急激に悪化した。過去の経験則が通じない状況に対して迅速に政策対応していくには、足元の経済・社会の動向を早期に把握する手段があると便利だ。

しかし、足元の経済状況をリアルタイムに把握することは非常に難しいという問題がある。通常、鉱工業生産指数などの経済指標（ハードデータと呼ばれる）の多くは数字が公表されるまでに1か月程度もしくはそれ以上のタイムラグがあり、足元の経済状況を迅速に把握することができない。そうした欠点を補うために、例えば日銀短観に代表される、比較的速報性の高いサーベイ調査により企業や家計の景況感を把握するという手法（これをソフトデータと呼ぶ）が用いられるが、あくまで実体経済を把握するための間接的な手法でしかないという限界がある。

そこで、ナウキャストイング（nowcasting）モデルが注目されている。そもそもナウキャストイングとは気象用語であり、数時間先といった近い将来の気象予報を行うことを指す。経済の文脈では、足元の経済状況を週次などの周期で早期に把握することを意味しており、ここ数年、海外を中心に直近に公表される四半期実質 GDP 成長率を早期に把握しようとするモデルの開発が進められてきた。特に今回のコロナ禍においては、急激な経済環境の変化を迅速に把握するため、こうしたナウキャストイングモデルに耳目が集まる場面が増えている²。

2. 現状のナウキャストイング

ナウキャストイングモデルの運用状況

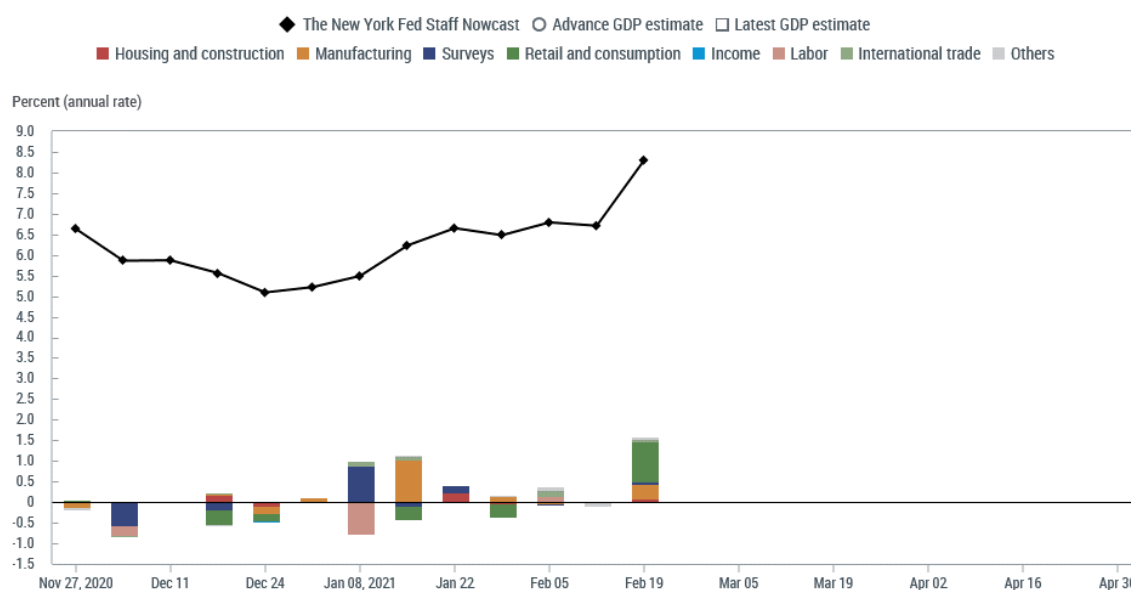
¹ 本シリーズのテーマをデータサイエンスとするのは、今回紹介するダイナミックファクターモデル（DFM）のようなマクロ経済分析に有効な新しい計量的手法やデータを幅広く捉えるためである。後述するように、機械学習やディープラーニングのようないわゆる AI 的な手法は、必ずしもマクロ経済分析に適さない側面があるにもかかわらず過度な期待が寄せられており、それが一般に大きな誤解を招いている。そうした認識から、本稿はより建設的な議論に資することを目的としている。

² マクロ経済予測の文脈では、数四半期先（短期）や数年先（中長期）の経済予測は通常の意味でフォーキャストイング（forecasting）と呼ばれる。さらに、マーケットが注目するような月次・四半期の経済指標を公表直前に予測する場合（超短期予測、例えば小峰 [1992] p57-70 を参照）は、既に起こったことを予測するという意味でバックキャストイング（backcasting）と呼ばれる。図表1の（注2）も参照されたい。

ナウキャストモデルは、Giannone et al. [2008]を嚆矢として、主に海外で研究されてきた。例えば、米国の中央銀行に相当するFRB傘下のNew York連銀³やAtlanta連銀⁴などでは、現在、独自のナウキャストモデルが運用されている⁵。

New York連銀のモデルを例にとると、月次や四半期の経済指標のうち比較的早く公表される指標から順次、統計モデルに取り込んでいくことで、注目する四半期実質GDP成長率の予測値を順次アップデートしていくという仕組みである。例えば、2021年第1四半期（1月～3月）の実質GDP成長率をナウキャストモデルで予測する場合、2021年1月時点ではサーベイデータのみを利用して、同年2月時点になると第1四半期のハードデータの一部（例えば同年1月分の雇用統計）が入手できるようになるので、それらを使って同年第1四半期の実質GDP成長率を予測する。次第に得られる経済データは増えてくるので、それらを加えてより精度の高い実質GDP成長率の予測ができるようになる。そして同年第1四半期の実質GDP成長率の速報値が得られる同年4月下旬に至るまで、足元で公表された多くの経済データを取り込んでいき、注目する四半期実質GDP成長率の予測値を、毎週、順次アップデートしていくのである（図表1）。

図表1 New York Fed Staff Nowcast による 2021 年第 1 四半期の実質 GDP 成長率の予測



(注1) 棒グラフは、直前の1週間で公表された各経済指標が実質GDP成長率予測のアップデートにどの程度寄与したのかを表している。

(注2) 厳密には、図表1（2021年第1四半期の実質GDP成長率の予測）の場合、2020年第4四半期以前の時点で予測する場合はforecasting、2021年第1四半期時点での予測はnowcasting、そして同年第2四半期に入ってから同指標の公表直前までの間に予測する場合はbackcastingと区別される。

(出所) New York Fed Staff Nowcastのウェブサイトより大和総研作成

³ New York Fed Staff Nowcast (<https://www.newyorkfed.org/research/policy/nowcast>)

⁴ GDPNow™ (<https://www.frbatlanta.org/cqer/research/gdpnow>)

⁵ 両者のナウキャストモデルの詳細について、New York連銀はBok et al. [2018]、Atlanta連銀はHiggins [2014]を参照されたい。日本を対象としたナウキャストモデルは、例えばBragoli [2017]やChikamatsu et al. [2018]がある。

ナウキャストモデルで主流となるダイナミックファクターモデル (DFM)

ナウキャストモデルは、通常、エコノミストが GDP の予測値を推計する方法と似ている。こうした実務で行われている予測手法を統計的に一貫した方法で行おうとするのがナウキャストモデルの問題意識となっている。その統計的手法にはいくつか種類があるが⁶、近年主流となっているのがダイナミックファクターモデル (DFM : Dynamic Factor Model) と呼ばれる手法である。

このモデルでは、マクロ経済指標は似たような変動を示すものが多いという統計的事実を踏まえて、その変動が何らかの少数の共通要因から生じていると考え、その共通要因を統計的に抽出する。そしてこの共通要因を景気変動とみなすのである。こうした手法では、扱うマクロ経済変数が多いことで生じる推計上の困難、いわゆる次元の呪い (curse of dimensionality) を回避できるという利点がある。

これに加えて、経済指標を活用する際に生じる別の問題を解決したことも、上記のモデルの開発が推し進められている理由である。例えば、実質 GDP 成長率の予測に必要な月次の経済指標の公表日がバラバラで、ある時点で予測をしようとするデータが出揃わずに欠損値が発生してしまうという問題や、月次や四半期などの周期が異なる経済指標が混在するといった問題だ。しかし近年、そうした問題を処理する手法 (例えば Bridge Equation Model やカルマンフィルターの適用等) が新たに活用されることで、上記のような状況下で生じる困難に対処できるようになった。

3. コロナ禍で浮上したナウキャストモデルの新たな課題とその解決

急激な経済環境の変化を追えないという課題

今回のコロナ禍で話題になっているのは、現在のナウキャストモデルが足元の急激な変化を十分に捉えきれなかったことだ。背景にあるのは、ナウキャストという名前が付いているものの、公表までにタイムラグのあるマクロ経済データを使っていることや、線形のモデルを前提としているため、新型コロナウイルス感染症による急激な経済悪化 (非線形的な動き) をモデルで追認できないことが挙げられている。これらがナウキャストモデルの新たな課題として浮上している。

ナウキャストモデルでビッグデータを活用

コロナ禍で浮かび上がったこうした課題解決に向けた動きとして、例えば Cristea[2020]では、電力消費データやトラック通行量データなどのリアルタイムで取得できるビッグデータを新たに用いて、より正確に欧州の景気動向を追認するためのモデルが研究されている。米国でも日次などの高頻度データを用いて週次単位で経済活動を捉える指標が開発されている (Lewis et al. [2020])。ところで、リアルタイム性を持つデータとしては伝統的に株価や為替、金利な

⁶ 例えば、ベイジアン VAR (BVAR : Bayesian vector autoregression) モデルがある。

どの金融データがあるが、ノイズが大きいことや経済指標（ハードデータやソフトデータ）を超える有益なマクロ経済情報を得にくく、ナウキャストモデルでの活用事例は多くない。しかし、今回のコロナ禍のように経済環境が急激に変化するような状況⁷や、予測期間に該当する経済指標がほとんど得られない予測の初期段階では、金融データもナウキャストモデルにおいて活用できる可能性がある。

欧州の事例で示したトラック通行量データのように、どのビッグデータが入手しやすいかは国そして分野により大きく異なり得るので、他国の事例がそのまま別の国でうまく機能するとは限らないが、ナウキャストモデルにビッグデータを取り入れることで、これまで以上にマクロ経済動向をリアルタイムにより高精度で把握できるようになるだろう。

4. 今後の展望

マクロ経済分野でもビッグデータや機械学習モデルの活用が進む可能性

これまでビッグデータや機械学習モデルのようなAI的手法は、マーケティングなどの分野で応用が進んできた。しかし、マクロ経済分野への応用は国内外でも遅れているのが現状である。その理由として、①ビッグデータを用いたマクロ経済予測のパフォーマンスが期待されたほど良くなかったこと、②機械学習のモデル構造がブラックボックスなために得られた数字の背景を人に説明できないという運用上の欠点があったこと、③大量のデータを使った分析を前提とする機械学習モデルではマクロ経済データは数が圧倒的に足りないこと、④マクロ経済分野で利用できるビッグデータの種類が限られること、などが挙げられる。

しかしながら、今後はマクロ経済分析全般においてもビッグデータや機械学習モデルの利用が進む可能性がある。

まず①に関しては、実は10年ほど前に、代表的なビッグデータの一つであるGoogle Trendsの検索データを使った予測モデルが注目された⁸が、期待されたほどの結果が得られずにその後の研究は下火となっていた。しかし、最近では労働市場や耐久財の需要予測などの特定分野では予測パフォーマンスが高いことや、従来のような単純な時系列モデルではなく機械学習の手法を使えば予測パフォーマンスが上がるということが認識されるようになり、ナウキャストモデルへの応用可能性が探られている(Woloszko[2020])。さらに、Google Trendsのようなビッグデータは、公的統計があまり整備されていない新興国や途上国におけるマクロ経済動向の把握手段⁹や、先進国でもサーベイデータの代替としてその有効性が主張されており、足元ではこうしたビッグデータをマクロ経済予測の文脈で再評価する動きが出てきている（この点は本シリーズにて今後取り上げるつもりである）。

次に②については、機械学習のようなモデルは非線形的な構造をしており、変数間の関係が非

⁷ Adrian et al. [2016]は、ダウンサイドリスクをモニタリングするのに金融データが有効だと指摘している。

⁸ 例えば、Choi and Varian[2012]などがある。

⁹ Carriere-Swallow and Labbe[2013]やNarita and Yin[2018]など。

常に複雑に入り組んでいるため、予測した結果を解釈するための背景説明が苦手であった。ところが、近年はどの要因がどれだけ予測に影響を与えるのかを説明できる機械学習モデル (Shapley Value を使った要因分解が行う研究) が開発されており (分かりやすい解説は Bolhuis and Rayner[2020] を参照されたい)、実際に上記の Woloszko[2020] では機械学習モデルを使った景気指標の変動要因を探るため、Shapley Value による要因分解を行っている。このように足元では機械学習を経済予測において運用する際に課題となっていた点を克服する動きが見られる。

一方、③に関しては、非線形な構造となる (より複雑な現象をモデル化できる) 機械学習モデルでは推計すべきパラメータの数が多いため、結果、必要とされるデータ数が非常に多くならざるを得ない。先程の Google Trends のようなビッグデータを使う場合には問題ないが、通常のマクロ経済データでは機械学習が要求するようなデータ数に足りないことも多く、そのままではうまく推計できない。そのため、次善策として利用するマクロ経済データの種類をできるだけ増やすか、DFM などのように別の方法で推計上の問題を回避するか、またはテキストデータを用いる¹⁰などの工夫が必要となる¹¹。

最後の④については、無料で利用できる公的統計を用いるマクロ経済分析とは異なり、ビッグデータの利用には様々な制約がある。一般的に個人情報や特定企業の情報を含むプライベートなデータであるため、そもそも利用できないかできても多額の費用が要求される、また、情報の取り扱いなどで法的な対応も必要になる。さらには、公的統計のような質の高いデータ (対象となる母集団を統計的に適切に表現したデータ [代表性を持つデータ] など) でないこともあり、以上のようなビッグデータ特有の問題がマクロ経済分析における活用で大きな障壁となっている。この点は一朝一夕では解決せず、官民学などが連携してビッグデータの利活用を促す制度整備などを行う必要があり、解決にはそれなりの時間を要するだろう。

しかしながら、海外の様々な研究が示すように、ビッグデータや機械学習などの AI 的手法がマクロ経済分析においても大きな期待を寄せられている点是不変である。今後、マクロ経済分析でも地に足の着いた形でデータサイエンスの活用が本格化するものと期待される。

以上

¹⁰ 実際、弊社が毎四半期に公表している「大和地域 AI (地域愛) インデックス」は、内閣府「景気ウォッチャー調査」のテキストデータをディープラーニングの手法で学習したモデルを使っている。

¹¹ ビッグデータや機械学習をマクロ経済分析 (特にナウキャスト) へ適用する際の統計的課題については、例えば Kapetanios and Papailias[2018] を参照されたい。

参考文献

- Adrian, T., N. Boyarchenko, and D. Giannone [2016], “Vulnerable growth,” *Staff Reports No. 794*, Federal Reserve Bank of New York, September 2016.
- Bok, B., Caratelli, D., Giannone, D., Sbordone, A. M. & Tambalotti, A. [2018], “Macroeconomic nowcasting and forecasting with big data”, *Annual Review of Economics* 10, 615-643.
- Bolhuis, M. A. and B. Rayner [2020], “Deus ex Machina? A Framework for Macro Forecasting with Machine Learning,” *IMF Working Paper* WP/20/45, International Monetary Fund.
- Bragoli, D. [2017], “Now-casting the Japanese economy,” *International Journal of Forecasting*, Volume 33, Issue 2, April-June 2017, 390-402
- Carriere-Swallow, Y. and F. Labbe [2013], “Nowcasting with Google Trends in an emerging market,” *Journal of Forecasting*, 32(4), 289-298.
- Chikamatsu, K., N. Hirakata, Y. Kido, and K. Otaka [2018], “Nowcasting Japanese GDPs,” *Bank of Japan Working Paper Series* No.18-E-18, November 2018, Bank of Japan.
- Choi, H. and H. Varian [2012], “Predicting the present with Google Trends,” *Economic Record* 88, 2-9.
- Cristea, R.-G. [2020], “Can Alternative Data Improve the Accuracy of Dynamic Factor Model Nowcasts? Evidence from the Euro Area,” *Cambridge Working Papers in Economics 20108*, Faculty of Economics, University of Cambridge.
- Ferrara, L. and A. Simoni [2020], “When are Google data useful to nowcast GDP? An approach via pre-selection and shrinkage,” Cornell University.
- Giannone, D., Reichlin, L. & Small, D. [2008], “Nowcasting: The real-time informational content of macroeconomic data,” *Journal of Monetary Economics* 55(4), 665-676.
- Higgins, P. [2014], “GDPNow: A Model for GDP “Nowcasting”,” *FEDERAL RESERVE BANK of ATLANTA WORKING PAPER SERIES* 2014-7, the Federal Reserve Bank of Atlanta.
- Kapetanios, G. and F. Papailias [2018], “Big Data & Macroeconomic Nowcasting: Methodological Review,” *ESCoE Discussion Paper 2018-12*, July 2018.
- Lewis, D., K. Mertens, and J. Stock [2020], “U.S. Economic Activity during the Early Weeks of the SARS-Cov-2 Outbreak,” *Staff Report No. 920*, Federal Reserve Bank of New York, April 2020.
- Narita, F. and R. Yin [2018], “Use of Google Trends’ data to narrow information gaps for low-income developing countries,” Technical Report WP/18/286, IMF Working Paper, 2018.
- Woloszko, N. [2020], “Tracking activity in real time with Google Trends,” *OECD Economics Department Working Papers* No. 1634, OECD.
- 小峰隆夫 [1992] 『基本ゼミナール 日本経済・景気 予測入門』 東洋経済新報社.