

# アセット・リターン予測 AI とマクロ経済理論の融合 —マルチタスク学習による正則化と識別—

Multi-task learning for asset prices and macro-economy inspired by theory-guided data science

塩野 剛志<sup>1\*</sup>

Takashi Shiono<sup>1</sup>

<sup>1</sup>クレディ・スイス証券株式会社 経済調査部

<sup>1</sup>Economic Research, Credit Suisse Securities (Japan) Limited

**Abstract:** 本研究では、資産価格リターンの予想を行う Residual Network (ResNet)を、マクロ経済の理論モデル (DSGE) とのマルチタスク学習により正則化する方法を提案する。経済理論が対象とする概念の一部は、現実の資産価格の決定メカニズムにも共有されている。したがって、マルチタスク学習によって共有因子を鮮明化することは汎化性能を改善する可能性がある。そこで、多数の経済金融時系列を入力とし、リターンを予測する ResNet の損失関数に対して、経済理論への整合性が制約となるようなペナルティを課す (正則化項の加算)、というマルチタスク学習を行った。実際、複数のハイパーパラメーターで検証した結果、マルチタスク化によって、テストデータへの汎化性能が改善する傾向が確認された。すなわち、多数の経済金融時系列を教師データとして受け取った AI は、リターン予想に資する特徴の抽出を、経済理論とある程度整合的な仕方で行い、それが過学習の予防となって汎化性能が向上した可能性がある。また、Layerwise Relevance Propagation を応用することで、マルチタスク共有中間層のうち、任意の経済ファクター (DSGE によって識別) と関連の深いユニットを特定し、それらが同時に資産リターンに与えた影響を抽出することで、AI 内部で形成されている経済ファクターと資産価格リターンの相関構造を記述することが可能となり、解釈性を高めることにもなる。

## 1.はじめに

日本においても、金融データ分析への AI (統計的機械学習) の応用が注目されるようになって数年が経ち、様々な実装が試みられてきただろう。当然のことながら、支配法則が確定しておらず、構造変化に開かれた動的開放システム (結果的に観測データは非定常) である経済金融現象においては、画像・音声認識、囲碁や将棋などと異なり、開発競争の末に圧倒的な性能のモデルがひとつに定まるということは起こりそうにない。どんな実装にも、その時々の相場・経済環境に応じて、実用上の一長一短が残り続けるだろう。本稿では、汎化性の高い予測を、なるべく解釈可能な枠組みで行う、という方向性を重視している。実際、AI の得意とするオルタナティブ・データからの特徴量抽出は、最終的にリターンなどの目標変数の外挿予測・説明に役立たなければ価値を持たないだろう。また、予測に用いることはせず、大量データの要約・解釈だけを追求す

るなら、むやみに複雑化せず、既存のマルチ・ファクターモデルやスタイル分析など線形モデルを使う方が望ましい場合も多い。つまり、常に圧倒的性能を保つ単一のモデル形式を得ることが難しいと思われる以上、AI 的手法を金融経済データの分析に利用するなら、できる限りよい汎化性能と人間による解釈の可能性をバランスを追求すること、が実用上の有益さに資すると考える。今回の分析は、そうした汎化性と解釈可能性 (識別) のバランスを目指した取り組みの一端である。具体的には、ResNet という比較的大きく、表現力の高い (それだけブラックボックス的でもある) 機構の学習に際して、マクロ経済の理論モデルとの整合性という制約を加えることで、中間層を規制し、正則化効果を得ると共に、解釈可能性の向上を図っている。以下では、まず、マルチタスク学習という、同時に複数のタスクに対して学習を進めることで正則化効果を得る枠組みの概略をまとめ (2 章)、続いて、本稿の提案手法である複数アセット・リターンとマクロ経済理論のマルチタスク学習モデルを紹介する (3 章)。その後、マルチタスク学習によって正則化効果が得られるか否かの検証実験を示し (4 章)、最後に、結果の解釈とそれに基づく議論を展開する (5 章)。

<sup>1\*</sup> 連絡先: クレディ・スイス証券株式会社 経済調査部, 〒106-6024 東京都港区六本木 1-6-1 泉ガーデンタワー26 階, Takashi.shiono@credit-suisse.com

## 2. マルチタスク学習

マルチタスク学習 (Multi-task learning; MTL) は、いくつかのタスクを同時に学習することで、タスク共通の要因を鮮明化し、汎化性能を改善する方法とされ、基本的な文献として [Caruana 97] が挙げられる。この試みが成功する条件 (仮定) は、異なるタスクを代表するデータ変動の説明因子の中に、互いに共有されるものが存在すること、である。

汎化性能が改善する理由としては、(1) 複数のタスクを同時に学習させることで、それらのタスクで共通に有用な特徴が選ばれ易くなる、(2) 各タスクに複数の局所解があるとき、複数のタスクに共通な局所解が選ばれ易くなる、(3) 共通する因子がある複数タスクがあれば、共通因子のノイズがタスク間のキャンセルで緩和される、(4) タスク間で学習の難易度に差があれば、困難なタスクは容易なタスクから情報を得てより簡単に学習可能、などの点が挙げられることが多い。

また、派生的な応用例としては、同じタスクを異なる損失関数を用いて異なるタスクのように学習し、損失関数毎の利点/欠点を補う枠組みや、既に大きなデータセットに対して学習済みのネットワークを新たな特化された目標に対して再学習させる (転移学習) などの枠組みがある。

本稿での取り組みに関して言えば、経済理論が対象とする概念の一部は、金融市場での資産価格決定メカニズムにも共有されているはずである (例えば、CAPM)。したがって、マルチタスク学習によって共有因子を鮮明化することは汎化性能を改善する可能性がある。また、今回のマルチタスク学習によって汎化性能の改善が確認されれば、経済ファンダメンタルズの理解が市場予想に対して無意味でないことの証左ともなり得る。

## 3. 提案手法

### 3.1. 概要

本稿の提案手法において、多数の経済金融時系列データを受け取った AI は、リターン予想に資する特徴の抽出を、経済理論とある程度整合的な仕方で行う。すなわち、資産価格リターンの予想を行う ResNet (Residual Network) を、マクロ経済の理論モデル (DSGE) とのマルチタスク学習により正則化する方法を提案する。

具体的には、多数の経済金融時系列を入力とし、リターンを予測する ResNet の損失関数 (メインタスク) に対して、同じ入力から DSGE モデルでマクロ経済変数へフィットする際の損失 (サブタスク)

を、一定の重みで加える、というマルチタスク学習を行う。直感的に言えば、AI は、リターンの予想に資する特徴の抽出を、経済理論とある程度整合的な仕方で行い、それが過学習の予防となって汎化性能が向上する可能性がある。

また、Layerwise Relevance Propagation を応用することで、マルチタスクの共有因子のうち、DSGE で意味特定 (識別) されたファクターに寄与したユニットを分別し、それらが同時に資産リターンに与えた影響を抽出することで、経済ファクターから資産価格リターンへの影響を算出することが可能となり、モデルの解釈性を高めることにもなる。

こうした取り組みは、データサイエンスと科学理論との融合を図る Theory-guided Data Science ([Karpatne et.al. 17]) の考え方に着想を得たものであり、ある意味では、AI の中間層に、理論的な考え方の枠組み、または思考のパターンを、実効的に「埋め込む」試みとも言える。

### 3.2. Residual Network

使用するデータを所与すれば、最良の性能は十分大きなモデルを適切に正則化できた時に得られるとされる ([Goodfellow et.al. 16])。今回はマルチタスクの共有部分に ResNet を使用しているが、同機構は CNN を多層化した比較的大きなモデルであり、表現力が高く様々な法則 (パターン) を捉えられる。その分、過学習を起こしやすく、正則化の意義は大きいはずである。

ResNet は、[He et.al. 2015]によって提案された畳み込み (Convolution) を多段に重ね、ショートカットを付けたアーキテクチャである。単純に畳み込みを多層にした場合、潜在的な表現力が上がるが、勾配消失で誤差逆伝播が進みづらく実用上の精度が高まらない問題があった。それに対し、ResNet では、ショートカット (ブロックの入力と出力の差分を学習) によって不要な変換をスキップし、教師信号がモデルの端まで伝播しやすくした点が革新的である。

なお、時系列データを時間軸方向に畳み込みすることは、時系列フィルタによる変数変換に相当する。この点、複数チャンネルの  $1 \times$  ラグウィンドウ幅の時系列入力に対する ResNet を定義すれば、事実上、不要ならスキップする多段の変数変換を施すことになると言えよう。例えば、「前年比」や「前期比」、「移動平均」などの変数変換も時系列フィルタの一種であり、ResNet を時系列データに用いることは、「前年比の前期差」、「3ヶ月移動平均の3ヶ月前比」などの適切な変数変換を AI が発見し、施してから回帰問題に持ち込むことに相当すると考えられる。

### 3.3. DSGE モデル

DSGE (Dynamic Stochastic General Equilibrium、動的確率一般均衡) モデルは、観測されるマクロ経済データの変動を、確率 (外生) 的に生じる「構造ショック」が、経済理論に基づく「遷移メカニズム」に沿って伝播した結果と見做す。マクロ経済学研究において一種のベンチマークとなっている標準モデルとして、Smets and Wouters [2007] (以下 DSGE-SW) がある。同モデルでは、マクロ経済には、1) 生産性、2) 家計選好、3) 外生需要、4) 投資効率、5) 物価マークアップ、6) 賃金マークアップ、7) 金融政策という 7 つの「構造ショック」が存在し、それを受けて、マクロ経済の状態を規定する次の 10 個の内生変数が決定されると考える (資本、労働投入、生産、消費、投資、実質資本利回り、実質資本価格、インフレ率、実質賃金、名目金利)。内生変数の遷移・決定メカニズムは、家計 (最終投資家)、労働仲介機関、中間財企業、最終財企業、資本形成部門、中央銀行と政府、という 7 部門の合理的経済主体の最適化行動から導出された構造方程式を対数線形近似したものとしてまとめられる。

内生変数の遷移メカニズムを規定する構造方程式のパラメータは、先行研究の結果を初期分布として採用しつつ、日本、米国、欧州、それぞれの経済データに合うように別途ベイズ推定しておく。こうした構造パラメータは各国のファンダメンタルな経済構造を反映しており、急激な時間変化はしないものと想定されている。なお、具体的な方程式体系と構造パラメータの推定結果については、補論にまとめている。

構造パラメータを所与とすれば、DSGE の対数線形化された構造方程式体系を内生変数について解くことで<sup>2</sup>、次式のような誘導形の線形状態空間システムが得られる。

$$U_t = FS_t,$$

$$S_{t+1} = PS_t + z_{t+1}$$

行列 F : 状態変数 S ⇒ 内生変数 U

行列 S : 構造ショック z を受けた状態変数 S の遷移

すなわち、DSGE モデルは、前期の状態変数を既知とした上で、今期の構造ショックを今期の経済データへと変換する行列に集約される。したがって、パラメタライズされた DSGE-SW の誘導形は、意味特定された構造ショックに相当する 7 次元の入力を受け取って経済データを出力する一種のニューラルネットワークのレイヤーと見做すことができる。後

<sup>2</sup> 実装には[Klein 00]の方法を用いた。

述するように、本稿では ResNet のサブタスクにおける出力層として日、米、欧それぞれの DSGE-SW の誘導形を実装している。

### 3.4. アーキテクチャと具体的な実装

改めて提案手法のアーキテクチャを見取り図で示せば図 1 のようになる。ResNet を共有ブロックとしたマルチタスク構造であり、メインタスクを複数資産リターンの 1 期先予想、サブタスクを DSGE レイヤーによるマクロ経済変数へのフィット、とした回帰問題である。入力データは、共有部の ResNet で変換された後、出力層に至る前に、タスク固有の中間層と教師なし中間層へと伝搬される。ここで、教師なし中間層は、入力データに含まれるが、どのタスクにも明確に寄与せず、かつ、ノイズとしてはキャンセルされなかったような残差的パターンを吸着しておく役割を期待したものである。

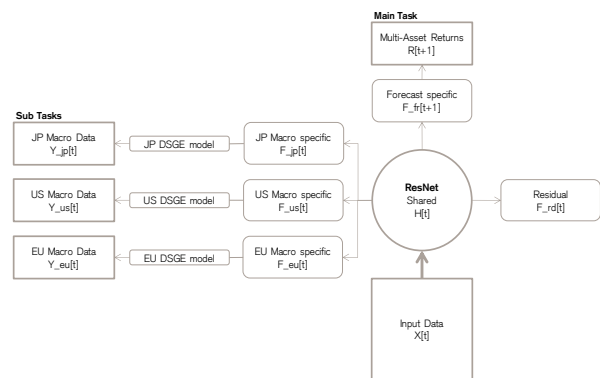


図 1 : 提案手法のアーキテクチャ

具体的な実装として、共有部の ResNet ブロックは、入力段階の Convolution1 層に加え、3 層のボトルネック型 ResNet ブロックを 4 つ重ねた 13 層構造。さらに、共有部からの出力は、各タスク固有の中間 1 層と各タスク固有の出力 1 層へと伝播し、全体の深さは計 15 層である<sup>3</sup>。

また、経済金融時系列データの 1 期分を  $X_t$ 、一度に入力するラグ期間の幅を  $W$  として  $([X_{t-W}, \dots, X_t])$ 、初期 Convolution 層に入力する 1 単位 (バッチサイズ  $b = 1$ ) の入力データは、たてよこ  $1 \times W$  サイズで、チャンネル数が  $\dim(X_t)$  のテンソル  $(b, 1, W, \dim X_t)$  とする。

この設計において、畳み込みカーネルは時間  $W$  方向を走査することになり、前述の通り、個々の時系列データはフィルタリングによって変数変換され

<sup>3</sup> 利用可能な計算資源、並びに、データ・サンプル数の限界を考慮して、ボトルネック型 ResNet ブロックの積層数は最小限にしており、ResNet 構造の実装としては小さい部類である。



ることになる。ResNet では、ブロック内でチャンネル次元の圧縮・拡張を繰り返すが、ここでは、チャンネル圧縮係数 ( $K$ ) を定義し、これが大きいほど最大チャンネルサイズ ( $\dim(X_t)/K$ ) が小さくなるように設計した。また、アーキテクチャに含まれる未知パラメータの学習は、アセット・リターン予想のバッチ平均二乗誤差に、日本、米国、欧州の DSGE による経済データ・フィットの二乗誤差を、ペナルティ・ウェイト ( $A$ ) を乗じた上で加算した損失関数を最小にするように進められる。

この際、入力データとリターンの教師データは週次を用いるが、マクロ経済データは四半期ベースの DSGE を通じてフィットさせるため、データ観測周期のマッチングをする必要がある。これには、モデル全体の訓練ミニバッチのサイズを 1 四半期分にする (バッチサイズ=約 13 週) ことで対応した。すなわち、ResNet を通じた構造ショックの週次理論値を DSGE レイヤーに入力すると、経済データの週次理論値が出力されるが、それをバッチ平均することで四半期ベースの理論値とし、マクロ経済データの四半期実績値との誤差を取る。そして、メインタスクである週次アセット・リターン予測誤差のバッチ平均 (四半期分) に、同期の経済データ誤差を加算したものを全体の損失関数として学習する。

以上の実装のもとで、探索するハイパー・パラメータは、畳込層のチャンネル圧縮係数 ( $K$ )、入力データのラグウィンドウ幅 ( $W$ )、DSGE ペナルティ・ウェイト ( $A$ )、の 3 つとする。なお、その他の設定として、最適化は Adam を使い、Epoch 数は Early Stopping により決定している。

### 3.5. 使用データ

今回使用した入力データは、次の国・地域について、ブルームバーグの<ECST>に掲載された「主要経済指標」(日、米、欧、英、豪、中、韓、台、印、タイ、インドネシア、ベトナム、ロシア、トルコ、南ア、ブラジル、メキシコ)、及び、その他グローバル金融指標、国際コモディティ市況 (CRB 内訳)、先進国セクター別株価指数、スマートベータ、クレジット指数、クロスカレンシー・スワップ等、計 2,036 系列である。

また、出力 (教師) データについて、メインタスクであるマルチ・アセットリターンは、クレディ・スイス計算による次の先進国株式と債券の先物インデックス (SP500 先物、DAX 先物、日経 225 先物、FTSE100 先物、米 10 年債先物、独 10 年債先物、日 10 年債先物、英 10 年債先物) と、為替レート (ドル円、ユーロ円、ポンド円) の週次変化率 (リターン)、計 11 系列である。

マクロ経済データについては、DSGE の内生変数と整合的になるように、日本、米国、欧州、それぞれの実質 GDP、消費、投資、労働投入、インフレ率、実質賃金、政策金利の 7 系列のデータを用いた。

なお、全サンプル期間は 2001 年 1 月第 1 週～2018 年 7 月最終週までである。

## 4. 正則化効果の検証実験

### 4.1. 性能評価の枠組み

ハイパー・パラメータの違いによる汎化性能の差を検証するため、性能評価の枠組みを設定する。今回は、複数テスト期間の平均パフォーマンスで評価する。なお、交差検証を用いる考え方もあるが、時系列データの予想問題においては、未来のデータで学習したパラメータを用いて過去を予想するテストになってしまうことから行わない。

具体的には、共通訓練期間を 2001～2007 年とし、その後 3 年ずつ訓練データを追加して 4 回バックテスト (テスト終点はいずれも 2018 年 3 月末週) を行う (テスト期間: ①2008～、②2011～、③2014～、④2017～)。このため、新たにデータを追加 (局面が変化) しても、高精度を維持したモデルが高評価となり、一時的によく当たっただけのモデルは評価されにくい。

また、パフォーマンスの評価尺度は、標準化 ( $z$  変換) した 11 資産のリターン (週次) の平均絶対予測誤差 (MAE) とする。

### 4.2. ランダムサーチによる検証

ハイパー・パラメータ探索にはランダムサーチを用いた。すなわち、ハイパー・パラメータの組み合わせを探索区間から乱数発生させて指定した。今回は 310 セットの乱数を発生させた。そのうち 40 は、経済データとのマルチタスク学習によるリターン予想問題の正則化効果の有無を図るのが目的のひとつであるため、ペナルティ・ウェイト ( $A$ ) を 0 に固定してコントロール・グループとした。

ランダムサーチを行い、前述の 4 回バックテストの平均絶対誤差 (MAE) でそれぞれのハイパー・パラメータの下でのモデルを評価した結果は、図 2、図 3、図 4、の通りとなった。すなわち、畳込層のチャンネル圧縮係数 ( $K$ ) は、0.3 と、2.5～3.0 範囲の 2 極で誤差が低下する傾向が見られた。また、入力データのラグ・ウィンドウ幅 ( $W$ ) は、10～20 の範囲で誤差が低下する傾向にあった。肝心の DSGE ペナルティ・ウェイト ( $A$ ) は、0.6～0.7 の範囲で誤差が低下する傾向が示された。

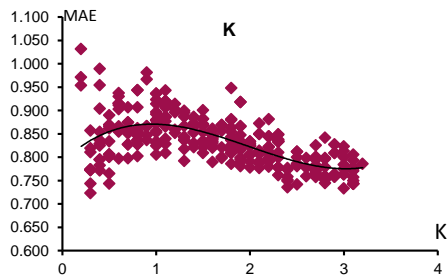


図 2 : K を指定した場合の MAE の分布

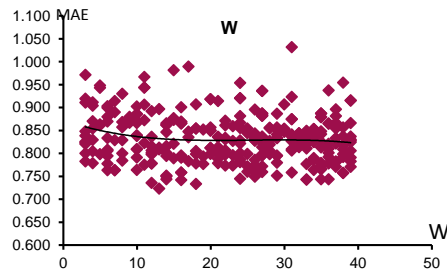


図 3 : W を指定した場合の MAE の分布

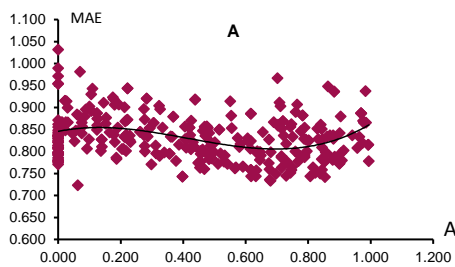


図 4 : A を指定した場合の MAE の分布

### 4.3.DSGE ペナルティによる正則化効果

改めて、DSGE ペナルティ・ウェイト ( $A$ ) が 0 のケース (他のパラメータはランダム) について、前述の 4 回バックテストでの  $z$  変換 11 資産平均絶対誤差 (MAE) の平均を計算すると、0.834 であった。他方、DSGE ペナルティ・ウェイト ( $A$ ) が 0.6 以上 0.7 以下の場合、同テスト値は 0.795 であった。すなわち、経済理論 (DSGE) ペナルティによって、確かにリターン予想の汎化性能は向上しており、正則化効果が確認された (図 5)。

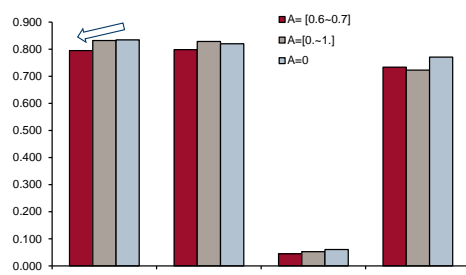


図 5 : A の領域別 MAE 分布の各種記述統計

ちなみに、「超過単位リスク当たりリターン」を評価尺度とした場合、最も優れていたスペックは  $K=2.3$ 、 $W=26$ 、 $A=0.57$  であった。この評価尺度は、AI による予測をシグナルとして利用し、対象の 11 資産について上昇予想ならロング (1)、下落予想なら 0 というポジションを取る単純なローテーション戦略のシャープレシオから、ベンチマークとして全 11 資産を全期間ロングするポジションのシャープレシオを引いた値である。もっとも、そのスペックでの実際の評価値 (超過シャープレシオの 4 回バックテスト平均) は年率換算で +0.14 と小幅であった (4 回テスト期間のシャープレシオの平均は AI シグナル運用で +0.87、ベンチマークで +0.73)。

## 5. 結果の解釈とモデルの副次的利用

### 5.1. 正則化効果の解釈

以上の検証の結果、DSGE を通じた経済データ説明とのマルチタスク化によって、複数資産のリターン予想における汎化性能が改善する傾向が示された。すなわち、多数の経済金融時系列を教師データとして受け取った AI は、リターン予想に資する特徴の抽出を、経済理論とある程度整合的な仕方で行い、それが過学習の予防となって汎化性能が向上した可能性がある。すなわち、深層学習の表現力の高いアーキテクチャでは、局在的 (時系列データで言えば短期的) にしか成立しないノイズのようなパターンであっても、訓練データにおける誤差低減への寄与が大きかったものなら重きを置いて学習してしまう。それに対して、経済理論をパラメータ学習の制約条件 (ガイド) として与えた場合、より普遍性のある特徴の抽出が喚起され、結果的に、まだ見ぬサンプルに対する予測力の向上に繋がったと考えられる。

また、マルチタスク学習による正則化が成功する条件は、異なるタスクを代表するデータ変動の説明因子の中に、互いに共有されるものが存在すること、であった。したがって、今回の結果は、経済ファンダメンタルズ理解が市場予想に対して無意味でないことの証左とも言える。資産価格の変動要因として、マクロ経済に関するファンダメンタル・リスクファクター (経済ファクター) が、確かに内在している。それを積極的に予想して投資機会とするか、または、予想を控え、複数の経済ファクターに分散投資して回避するかは戦略次第だが、いずれにせよ、AI であろうとなかろうと、マクロ経済ファクターへの理解が運用パフォーマンス向上に資する可能性が示唆されたと言える。

## 5.2. マルチタスクによる学習へのガイド

直感的に言えば、本稿提案手法において、AI は入力に対してどう反応すべきかのガイドを与えられた上で予想を作成している。つまり、既に普遍性や有用性が確認されている考え方の枠組み（理論）や思考パターン（経験則）を AI の中間層へ「埋め込んで」いる」との表現も可能であろう。

こうした「中間層形成に対する思考のガイド」が経済理論である必然性はなく、また、予測対象がマクロの資産インデックスである必要もない。例えば、決算情報やテキストデータを入力とし、「典型的な運用スタイル」や「著名投資家の行動履歴」をガイドにした上で、個別株のストック・ピックを行うような判別問題のアーキテクチャを構築してもよいだろう。

## 5.2. 経済ファクターとリターンの関係を識別

続いて、経済理論をガイドとしたマルチタスク学習の副次的な利点として、経済ファクターと資産リターンとの関係性のある程度識別することができる。繰り返しとなるが、マルチタスクによる正則化効果があったということは、リターン予想と経済ファンダメンタルズに共有因子あることを意味する。

そのため、LRP (Layerwise Relevance Propagation)<sup>4</sup>を用いて、マルチタスク共有因子のうち、それぞれの経済ファクター (=DSGE 構造ショック) の出力に寄与したユニットを絶対値でランキングする (今回は上位 5%のみ指定)。これによって、その経済ファクターと関連の深い共有部ユニットがどれかを識別することができる。ひとたび共有部における任意の経済ファクターとの関連ユニットが特定できれば、入力データに対して、その関連ユニットのみを使用して予想リターンを出力することで、予想リターンのうち、ある経済ファクターとの関連が特定された成分を生成することができる (図 6)。

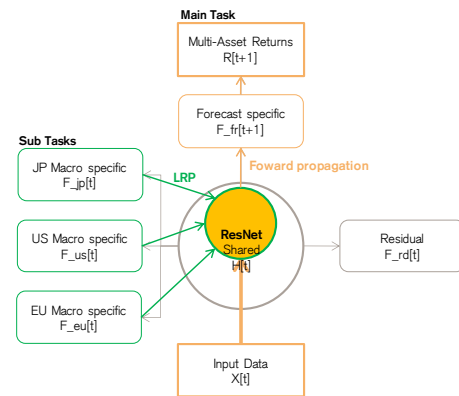


図 6：サブタスクに関連するユニットの特定と使用

なお、こうした識別の作業を行う際には、直近までのデータを全て訓練データとして学習させ、出力は全てインサンプル理論値となっている。また、ハイパー・パラメータは  $K=0.3$ ,  $W=15$ ,  $A=0.6$  を選択している。計量経済学では全サンプルを使用した線形回帰を使って変数間の関連識別を行うことが多いが、今回の作業は同様の発想を大規模データとニューラルネットの下で実装する工夫であると言える。

実例として、入力データを受け取った AI の共有中間層 (脳のようなもの) のうち、主として米国の家計リスク選好ファクターに関連することが分かっているユニット (脳の領域) のみを活性化して出力した米国株価リターンの成分は、図 7 のように示される。

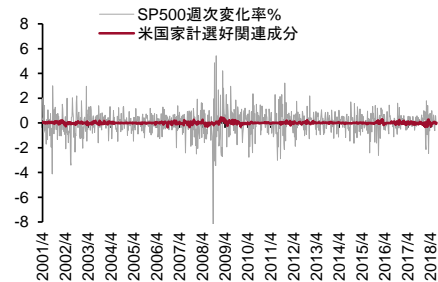


図 7：米国家計選好関連ユニットによる米株リターンの理論値

また、このように識別した各経済ファクターの各資産リターンへの影響成分 (識別ユニット出力) と、実際の各資産リターンの実績値との相関を示したのが図 8 である (2002~2017 年)。全ユニットの上位 5% という限られた情報のみを活性化して得た成分であるため、リターン実績全体の動きと比べた相関係数の値は小さくなりがちだが、資産クラス毎の各ファクター間の相対順位には意味がありそうだ。例えば、今回の AI は、訓練期間の学習の結果、米国株価の変動予想について、「米国家計のリスク選好」ファクターと、「米国の物価マークアップ」ファク

<sup>4</sup> LRP は出力結果に対する入力データの寄与度を、推定済みのニューラルネットワークを出力層から逆に遡及計算していく手法である。詳しくは [Binder et al. 16] を参照されたい。

ターが特に正相関していると判断していた。

	JP				US				EU			
	zA	zB	zI	zP	zA	zB	zI	zP	zA	zB	zI	zP
SP500先物	0.01	0.02	0.03	-0.03	0.01	0.03	0.08	-0.01	0.08	0.05	0.02	0.08
日経225先物	0.01	0.00	0.03	0.04	0.03	0.00	0.04	0.00	0.01	0.01	0.04	0.02
DAX先物	-0.04	-0.02	0.00	-0.02	0.00	-0.04	-0.04	-0.04	0.01	0.00	0.03	0.00
FTSE100先物	0.02	0.03	0.03	0.04	0.01	0.03	0.04	0.01	0.02	0.02	0.03	0.01
米10年債先物	0.00	0.03	0.01	0.05	0.04	0.02	0.05	0.03	0.04	0.05	0.04	0.00
日10年債先物	0.07	0.05	0.09	0.05	0.06	0.07	0.02	-0.01	0.08	0.03	0.02	0.06
独10年債先物	0.07	0.02	0.03	0.07	0.05	0.01	0.07	0.05	0.03	0.07	0.04	0.08
英10年債先物	0.03	0.07	0.00	-0.02	0.03	0.01	0.04	0.02	0.04	0.05	-0.03	-0.02
ドル円	0.05	0.05	0.03	0.03	0.08	0.07	0.07	0.06	0.05	0.07	0.02	0.05
ユーロ円	-0.01	-0.02	-0.01	0.01	0.00	0.00	-0.03	-0.04	0.00	-0.02	-0.01	-0.03
ポンド円	0.04	0.03	0.01	0.05	0.00	0.01	-0.03	-0.03	0.00	0.03	0.01	0.00
生産性 (zA)												
家計選好 (zB)												
外生需要 (zI)												
投資効率 (zP)												
物価マークアップ (zW)												
賃金マークアップ (zX)												
金融政策 (zR)												

図 8：識別ユニット出力と各資産リターンの実績値との相関構造

また、AI のアウトオブ・サンプル予想が外れた場合に、原因究明が求められることがある。その際は、ある資産クラスについて、アウトオブ・サンプル出力における各経済ファクターとの相関構造と、改めて全データで AI を訓練し直した場合のインサンプルでの相関構造との差を見ることで、過去のデータからは獲得できていなかったサプライズなパターンがどの経済ファクターについて生じたのか、までは識別することができよう（例えば、訓練期間では米国株に米国生産性はあまり効いていなかったのに、テスト期間では予想外に強く効き始めてしまったために予測精度が低下した、などと記述が可能）。

さらに、こうしたインサンプル推計による識別ユニット出力を使って、過去局面毎に、経済ファクターと資産リターンとの関係性がどう変化してきたかを確認することもできる。例えば、図 8 は、米国家計選好の改善シグナルで米株をロング、悪化シグナルでショートした場合の累積リターン（＝米国家計選好ファクター・プレミアム）と、米国物価マークアップの上昇シグナルで米株をロング、悪化シグナルでショートした場合の累積リターン（＝米国物価マークアップ・ファクター・プレミアム）を示したものである。累積リターンが拡大している期間は、米国家計のリスク許容度が改善すれば株価が上昇するという理論的に期待される関係が正常に機能していた期間であり、累積リターンが縮小している期間はその法則が機能しなくなった期間（ある種の異常性の検知）と解釈できる。



図 8：経済ファクター・プレミアムの例

### 5.3. ナウキャスト

最後に、本稿提案アーキテクチャの別の副次的な利用として、サブタスクの出力による経済データのナウキャストが可能である。同 AI のアウトオブ・サンプル期間におけるサブタスク出力は、各時点で利用可能な入力データを用いて、その時点が含まれる四半期のマクロ経済データの状態を週次で推定したものととなる。これは、最新時点での入力データで変動が分かっている金融データや高頻度の先行指標などから、発表ラグのある GDP を中心とした出力側のマクロ経済データを推定したナウキャスト値としてそのまま利用可能である（図表 9、10）。これは、マーケットが大量情報を消化して資産価格を変化させるプロセスを追うことで、そこで想定されているマクロ経済要因の変化を察知するプロセスと言える。

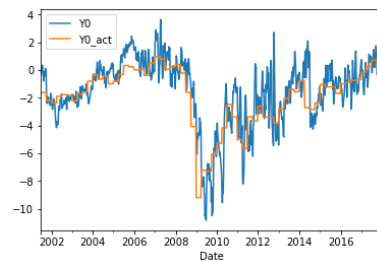


図 9：日本実質 GDP（対数・線形トレンド除去%）：内挿理論値

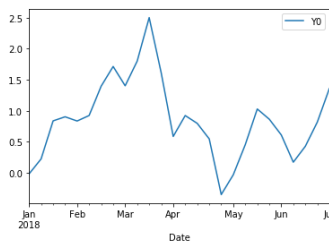


図 10：同・外挿予想値

### 参考文献

[Binder et al. 16] Binder, A. "Layer-wise Relevance Propagation for Neural Networks with Local Renormalization Layers", arXiv:1604.00825 (2016).

[Caruana 97] Caruana, R. "Multitask Learning", Machine Learning, vol.28 (1997).

[Goodfellow et al. 16] Goodfellow, I. et al. "Deep Learning", MIT Press (2016).

[He et al. 2015] Kaiming, H. et al. "Deep Residual Learning for Image Recognition," arXiv:1512.03385 (2015).

[Karpatne et al. 17] Karpatne, A. et al. "Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data," IEEE Transactions on Knowledge and Data Engineering, vol. 29, no. 10,



pp. 2318-2331, (2017). doi: 10.1109/TKDE.2017.2720168

[Klein 00] Klein, P. "Using the generalized Schur form to solve a multivariate linear rational expectations model", Journal of Economic Dynamics and Control vol. 24 No.10, pp.1405-1423 (2000).

[和泉 03] 和泉 潔: 人工市場—市場分析の複雑系アプローチ, 森北出版 (2003).

https://www.morikita.co.jp/books/book/2239

## 補論

### (1) DSGE-SW の構造形方程式 (線形近似)

[SW1] 資本成長

$$K_t = (1 - \tau)K_{t-1} + \tau I_t + \tau(1 + \beta)S^m \epsilon_t^I,$$

[SW2] 労働需要

$$L_t = -w_t + \left(1 + \frac{1}{\psi}\right)r_t^k + K_{t-1},$$

[SW3] 資源制約

$$Y_t = C_t + I_t + \frac{r^k K_t}{\psi} + \epsilon_t^C,$$

[SW4] 生産関数

$$Y_t = \phi \epsilon_t^a + \phi \alpha K_{t-1} + \frac{\phi \alpha}{\psi} r_t^k + \phi(1 - \alpha)L_t,$$

[SW5] 一般化テイラールール

$$R_t = \rho R_{t-1} + (1 - \rho)[r_{\pi 1} \pi_t + r_{y 1} Y_t + r_{\pi 2} \pi_{t-1} + r_{y 2} Y_{t-1}] + \epsilon_t^r,$$

[SW6] 消費遷移式

$$C_t = \frac{h}{1+h} C_{t-1} + \frac{1}{1+h} E_t[C_{t+1}] - \frac{1-h}{(1+h)\sigma_c} (R_t - E_t[\pi_{t+1}]) + \epsilon_t^b,$$

[SW7] 投資遷移式

$$I_t = \frac{1}{1+\beta} I_{t-1} + \frac{\beta}{1+\beta} E_t[I_{t+1}] + \frac{1}{(1+\beta)S^m} q_t + \epsilon_t^I,$$

[SW8] 資本価格決定式

$$q_t = -(R_t - E_t[\pi_{t+1}]) + \frac{1 - \tau}{1 - \tau + r^k} E_t[q_{t+1}] + \frac{r^k}{1 - \tau + r^k} E_t[r_{t+1}^k],$$

[SW9] 物価フィリップス・カーブ

$$\begin{aligned} \pi_t &= \frac{\beta}{1 + \beta t_p} E_t[\pi_{t+1}] + \frac{t_p}{1 + \beta t_p} \pi_{t-1} \\ &\quad + \frac{(1 - \beta \xi_p)(1 - \xi_p)}{(1 + \beta t_p) \xi_p} (\alpha r_t^k + (1 - \alpha)w_t - \epsilon_t^a) \\ &\quad + \epsilon_t^p, \end{aligned}$$

[SW10] 賃金フィリップス・カーブ

$$\begin{aligned} w_t &= \frac{\beta}{1 + \beta} E_t[w_{t+1}] + \frac{1}{1 + \beta} w_{t-1} + \frac{\beta}{1 + \beta} E_t[\pi_{t+1}] + \frac{1 + \beta t_w}{1 + \beta} \pi_t \\ &\quad + \frac{t_w}{1 + \beta} \pi_{t-1} \\ &\quad - \frac{(1 - \beta \xi_w)(1 - \xi_w)}{(1 + \beta) \left(1 + v_l \frac{1 + \lambda_w}{\lambda_w}\right) \xi_w} (w_t - v_l L_t) \\ &\quad - \frac{\sigma_c}{1 - h} (C_t - h C_{t-1}) + \epsilon_t^w, \end{aligned}$$

[AR1] 生産性ショックの AR(1)過程

$$\epsilon_t^a = \rho_a \epsilon_{t-1}^a + z_t^a,$$

[AR 2] 選好ショックの AR(1)過程

$$\epsilon_t^b = \rho_b \epsilon_{t-1}^b + z_t^b,$$

[AR 3] 外生需要ショックの AR(1)過程

$$\epsilon_t^c = \rho_c \epsilon_{t-1}^c + z_t^c,$$

[AR 4] 投資効率ショックの AR(1)過程

$$\epsilon_t^I = \rho_I \epsilon_{t-1}^I + z_t^I,$$

[AR 5] 物価マークアップ・ショックの AR(1)過程

$$\epsilon_t^p = \rho_p \epsilon_{t-1}^p + z_t^p,$$

[AR 6] 賃金マークアップ・ショックの AR(1)過程

$$\epsilon_t^w = \rho_w \epsilon_{t-1}^w + z_t^w,$$

[AR 7] 金利環境 (裁量金融政策) ショックの等号

$$\epsilon_t^r = z_t^r,$$

内生変数		外生 AR (1) 過程		構造ショック	
$K_t$	資本	$\epsilon_t^a$	生産性	$z_t^a$	生産性ショック
$L_t$	労働投入	$\epsilon_t^b$	選好	$z_t^b$	選好ショック
$Y_t$	生産	$\epsilon_t^c$	外生需要	$z_t^c$	外生需要ショック
$r_t^k$	実質資本利回り	$\epsilon_t^I$	投資効率	$z_t^I$	投資効率ショック
$R_t$	名目金利	$\epsilon_t^p$	物価マークアップ	$z_t^p$	物価マークアップ・ショック
$C_t$	消費	$\epsilon_t^w$	賃金マークアップ	$z_t^w$	賃金マークアップ・ショック
$I_t$	投資	$\epsilon_t^r$	金利環境	$z_t^r$	金利環境ショック
$q_t$	実質資本価格				
$\pi_t$	インフレ率				
$w_t$	実質賃金				

### (2) DSGE-SW の構造パラメータ推定結果

パラメータ 役割	事前分布	推定結果 (事後分布平均)			
		形状	平均	標準偏差	95% 区間
<b>経済構造パラメータ</b>					
$\alpha$	資本係数 (分配率)	外生	-	0.30	0.30
$\beta$	主観的時間割引率	外生	-	0.99	0.99
$\tau$	資本減耗率	外生	-	0.03	0.03
$\lambda_w$	賃金マークアップ係数	外生	-	0.30	0.30
$\psi$	資本稼働コスト	ベータ	0.35	0.08	0.38
$t_p$	物価のインデジゼーション度	ベータ	0.26	0.15	0.34
$t_w$	賃金のインデジゼーション度	ベータ	0.22	0.15	0.29
$\xi_p$	Calvo の価格硬直性	ベータ	0.84	0.1	0.86
$\xi_w$	Calvo の賃金硬直性	ベータ	0.85	0.15	0.92
$v_l$	労働の CRRA 係数	ガンマ	2	0.15	2.11
$\sigma_c$	消費の CRRA 係数	ガンマ	1.7	0.15	1.81
$h$	慣性形成度	ベータ	0.69	0.1	0.68
$\varphi$	生産の固定費用	ベータ	0.45	0.1	0.52
$S^m$	資本調製コスト係数	正規	5.3	1	5.30
<b>政策反応パラメータ</b>					
$f_{\pi 1}$	現在インフレへのテイラールール係数	ガンマ	2.2	0.1	2.31
$f_{y 1}$	現在生産へのテイラールール係数	ベータ	0.35	0.1	0.40
$f_{\pi 2}$	過去インフレへのテイラールール係数	正規	-0.22	0.1	-0.22
$f_{y 2}$	過去生産へのテイラールール係数	正規	-0.08	0.1	-0.03
$\rho$	名目金利の自己回帰係数	ベータ	0.87	0.1	0.92
<b>外生過程パラメータ</b>					
$\rho_a$	生産性 AR(1)過程の自己回帰係数	ベータ	0.91	0.1	0.92
$\rho_b$	選好 AR(1)過程の自己回帰係数	ベータ	0.77	0.1	0.95
$\rho_c$	外生需要 AR(1)過程の自己回帰係数	ベータ	0.9	0.1	0.76
$\rho_I$	投資効率 AR(1)過程の自己回帰係数	ベータ	0.71	0.1	0.94
$\rho_p$	物価マークアップ AR(1)過程の自己回帰係数	ベータ	0.82	0.1	0.77
$\rho_w$	賃金マークアップ AR(1)過程の自己回帰係数	ベータ	0.52	0.1	0.83
<b>構造ショックの標準偏差</b>					
$\sigma_a$	生産性ショックの標準偏差	逆ガンマ	0.1	0.5	0.82
$\sigma_b$	選好ショックの標準偏差	逆ガンマ	0.1	0.5	0.83
$\sigma_c$	外生需要ショックの標準偏差	逆ガンマ	0.1	0.5	0.79
$\sigma_I$	投資効率ショックの標準偏差	逆ガンマ	0.1	0.5	0.81
$\sigma_p$	物価マークアップ・ショックの標準偏差	逆ガンマ	0.1	0.5	0.79
$\sigma_w$	賃金マークアップ・ショックの標準偏差	逆ガンマ	0.1	0.5	0.81
$\sigma_r$	裁量金融政策ショックの標準偏差	逆ガンマ	0.1	0.5	0.81