



RIETI Discussion Paper Series 21-J-049

## 機械学習手法を用いた勘定科目レベルの異常検知

宇宿 哲平

あずさ監査法人

近藤 聡

あずさ監査法人

白木 研吾

あずさ監査法人

眞田 貴央

あずさ監査法人

須崎 公介

あずさ監査法人

宮川 大介

一橋大学



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所

<https://www.rieti.go.jp/jp/>

## 機械学習手法を用いた勘定科目レベルの異常検知\*

宇宿 哲平（あずさ監査法人）、近藤 聡（あずさ監査法人）、白木 研吾（あずさ監査法人）、  
眞田 貴央（あずさ監査法人）、須崎 公介（あずさ監査法人）、宮川 大介（一橋大学）

### 要 旨

本研究は、本邦上場企業の不正会計イベントを対象として、勘定科目レベルでの異常検知（不正検知）を行ったものである。教師無し学習手法の一種であり、変数間の偏相関関係の推定を、次元圧縮を行いつつ実施する「スパースモデリング手法」（graphical lasso）を用いた分析手法を実装するとともに、out of sampleでの検知能力をテストしたところ、実務的な観点から十分な検知精度を達成したほか、スパース化による検知精度の改善も図られていることが確認された。

キーワード：不正会計、勘定科目、スパースモデリング

JEL classification: M42, C53, C14

RIETI ディスカッション・ペーパーは、専門論文の形式でまとめられた研究成果を公開し、活発な議論を喚起することを目的としています。論文に述べられている見解は執筆者個人の責任で発表するものであり、所属する組織及び（独）経済産業研究所としての見解を示すものではありません。

\*本稿は、国立大学法人一橋大学と有限責任あずさ監査法人との共同研究契約に基づき実施された研究プロジェクト及び経済産業研究所「企業金融・企業行動ダイナミクス研究会」（プログラム：産業フロンティア）において実施された研究プロジェクトの成果物である。

## 1. はじめに

企業による不正会計（いわゆる粉飾決算）の検知と予測は、その実務的な重要性を背景として実証会計分野における重要な研究テーマとして位置づけられてきた。国内外における近年の不正会計の事例は、不正の発覚に伴って有形無形の制裁を受ける事業会社の側のみならず、監査業務を担う監査法人側における責任を厳しく問う内容となっている。監査法人が従前から担ってきた労働集約的な監査業務を補完する形で、より高精度かつ実装可能な不正検知・予測の仕組みづくりに関するニーズが高まっているといえるだろう。

これまで、観察可能な情報を用いて現時点若しくは将来時点における不正会計の発生を検知もしくは予測する試みは、主として企業レベルのデータを用いた実証研究の形態で蓄積されてきた（Dechow et al. 1996, 2010, 2011; Song et al. 2016）。伝統的には、企業が不正にコミットする何らかの理論的なメカニズムを想定したうえで選択された説明変数を用いた古典的なパラメトリックモデルによる不正の検知・予測が多く行われてきたが、近年では利用可能な企業属性変数の高次元化に対応する形で機械学習手法を全面的に用いたモデル構築も進んでいる（Dechow et al. 1996, 2010, 2011; Song et al. 2016）。

企業レベルの不正会計検知・予測の試みは、実務的に見ても有用な知見を提供するものである。例えば、特定の企業に関して現時点における不正の蓋然性が高いという情報が当該企業を担当する会計士に伝えられたとする。こうした情報を監査法人が有効に活用しようとした場合、例えば、当該企業に関与する会計士の人数を増強するなどの労働資源の望ましい再配分が行われる可能性もある。

こうした監査法人全体の資源配分という観点から見た有用性の一方で、個々の監査担当者の立場からは、企業レベルの不正検知・予測のみではその実務的な要請に応え

られない場合もあるだろう。例えば、企業レベルで過去の不正事例との類似度を測る教師あり学習を前提とすると、例えば、新たなビジネスにおいてこれまでと全く異なる不正が行われた場合に検知・予測できない可能性がある。また、担当企業に関する不正の蓋然性が高いとしても、会計情報のどの部分へより多くの資源を投入して監査業務を行えばよいかは明らかにはならない。

こうした問題意識を踏まえて、本研究では、企業レベルの不正会計検知・予測を主たる目的とした先行研究を補完する形で、勘定科目レベルの不正検知に向けたモデル構築を行う。具体的には、各企業の勘定科目レベルで計測された不正に関するスコアを提供する枠組みを構築し、その上で実際のデータを用いてその枠組みの有効性を実証的に検討する。

本研究で取り扱うモデルは、異常検知を目的として工学分野で研究の蓄積が進んでいるものである。その基本的なアイデアは、勘定科目間の共分散構造に関する推定結果を前提として、関心のある科目について、当該科目以外の科目の観測値で条件付けた分布を求め、当該科目の実測値がこの条件付き分布を前提とした場合にどの程度「異常」な値を示しているかを計測するというものである。このとき問題となるのは、分析に含めるべき勘定科目が膨大な数に及ぶという点であり、上記の条件付き分布を求める際の条件に含まれる変数の次元を圧縮（スパース化）する実務的な必要性が生じる。既に先行研究の一部では、VAR 形式の LASSO 推定でこうした問題を解決したものが認められるが、本稿では、異常度計測に関してより理論的に厳密な方法である **graphical lasso** (Friedman et al. 2008)を用いることで、変数間の偏相関関係の推定を、次元圧縮を行いつつ実施する。

本研究の結果は以下の通りである。第一に、**graphical lasso** による科目間共分散構造の推定結果を基にして構築した各科目の条件付き分布から各科目の異常度を求めるモデルを構築したところ、**out of sample** での AUC が 0.75 から 0.8 程度となり、実務的に

も有効な精度が実現された。第二に、こうした精度を実現するにあたって、上記のスパース化が一定の役割を果たしていることを確認した。

次節以降の構成は以下の通りである。第二節では、本研究で使用するデータの概要を説明する。第三節では、本研究で構築したモデルの概要を説明する。第四節では、予測精度の検証結果を示すと共に、第五節で今後の研究に向けたディスカッションを行う。第六節では本稿のまとめを示す。

## 2. データ

本研究では、以下の二種類のデータを用いる。第一に、上場企業の財務諸表情報であり、この情報は貸借対照表 (BS)、損益計算書 (PL)、キャッシュフロー計算書 (CF) から構成される。第二に、これらの上場企業に関する過去の会計不正イベントフラグである。このフラグ情報は、不正の種別ごとに、売上過大、費用過小、資産過大の三つの区分と共に記録されている。

モデルの構築と精度検証において用いる変数群は、第一のデータに含まれる企業毎に計測された 399 種類の科目情報である。表 1 は、これらの科目の数を示したものである<sup>1</sup>。

表 1 : モデル作成に用いる科目数

BS	PL	CF
226 科目	159 科目	14 科目

個々の企業について複数の年度に亘って記録されたこれらの情報の中には、本来報

---

<sup>1</sup> 後述の通り、分析に当たってのデータ前処理の過程でこれらの変数からの絞り込みが行われるため、実際に使用される科目数は表 1 記載のものよりも少なくなる。

告されるべき数値と異なる内容が記録されていたことから訂正を要するものが含まれており、本研究で用いるデータにはこうした訂正履歴が含まれている。最新時点で取得したデータにはこれらの訂正が完了した数値が記録されているため、本研究では、訂正履歴の有無に応じて、訂正歴を有するデータについては訂正前の数値を復元し、訂正履歴がない科目については最新時点で取得した数値をそのまま用いることとする。こうした処理により、財務諸表の各報告時点における内容を分析に用いることが可能となる。

既述の不正イベントに対応したフラグ情報については、宇宿ほか(2019)などの先行研究に従い、有価証券報告書の虚偽記載で金融庁の課徴金勧告の対象となったケース及び証券取引委員会の告発の対象となったケースを参照して構築した。この時、各ケースの具体的な中身を精査したうえで、売上過大、費用過小、資産過大の三種類の不正フラグを準備した。

上記の処理の結果として、2005年1月から2018年10月までの期間における5663社、57679社×年を対象としたデータセット（各社×年毎に科目の数値情報が記録されたパネルデータセット）が構築された。

### 3. モデル

本研究では、Banerjee et al. (2006)によって基本的な問題の定式化が行われ、Friedman et al. (2008)でその効率的な解法が提案された graphical lasso 手法を用いる。同手法は、時系列の記録を持つ複数の変数系列間における偏相関関係について、高速かつ安定的なスパース化を実行しながら共分散構造に関する情報を抽出するものである。こうした共分散構造に関する情報は、当該の構造を仮定した場合に生じる各変数の条件付き分布を求める際に参照することが可能であり、実際に観察された変数の異常度を計測す

る目的で利用されている（例：井手 2017; 井手剛, 杉山将 2015; Idé et al. 2009）。本研究では、主として工学分野で応用されてきたこの **graphical lasso** 手法を、企業の財務諸表を構成する科目間の共分散構造の推定に応用することで、観測された各科目の異常度を計測する。

以下では、まず、前節で概観した財務データを **graphical lasso** 手法で利用するための前処理工程を説明する。第一に、共分散構造の推定において参照する変数系列の絞り込みを行う。この処理の目的は、データの生成過程で機械的な理由によって極端に高い相関関係にある科目を分析から落とすことにある。具体的には、99%以上のセルが一致する科目同士は一方のみを残す処理を行う。また、科目の欠損有無がモデル構築後の精度評価において目的変数として参照する「訂正の有無」と独立していない科目を共分散構造の推定において含めると、極めて高い予測精度が実現してしまうため、科目の欠損有無と訂正の有無の独立性検定（カイ二乗検定）における  $p$  値が 0.01 未満の科目は削除した。更に、変数の欠損割合が 99%以上の科目についても削除した。

第二に、欠損値の補完処理を行う。まず、原則として、欠損しているセルは、同じ企業の当該科目の平均値（複数年度の平均）で補填する。次に、同じ企業の当該科目がすべて欠損している場合は、同じ業種の平均値で補填する。最後に、同じ業種の当該科目がすべて欠損している場合は、すべての企業の平均値で補填する。

第三に、共分散構造の推定に当たって参照するレコードの絞り込みを行う。まず、売上過大を対象とする場合、売上以外の不正（費用過小・資産過大）のレコードを削除する。費用過小・資産過大の場合も同様に、ほかの不正のレコードを削除する処理を行う。こうした処理を行う目的は、共分散構造の推定に当たって不正の含まれない「**healthy データ**」のみを参照することで、正常な状態を描写する **graphical model** を構築することにある。こうした処理を行わない場合、不正を行っているレコードが学習用データに含まれる結果、共分散構造の推定後に行う異常度の計測が適切に行われな

い可能性がある。次に、売上過大を対象とする場合、売上高の修正率が0%以下のレコードを削除する。費用過小・資産過大の場合も同様の処置を行う。最後に、今回の分析では、科目間の共分散構造に一定の共通点がある企業群に分析を絞る趣旨から、製造業に属する企業のレコードのみを用いることとする。表2は、これらの処理を行った上でモデル構築用に利用可能な変数の総数とレコード数をまとめたものである。

表2：データ加工終了時のデータ概要

	変数の総数	レコード数
売上過大モデル	176	22,476
費用過小モデル	209	22,462
資産過大モデル	162	22,500

第四に、データを企業×科目の粒度で標準化する。Graphical lassoに限らず正則化項を含むモデルを用いた変数の次元圧縮を行う場合には、各変数に付随して推定されるパラメータのサイズに着目した処理が行われるため、推定に用いられるデータが比較可能な振る舞いを示している必要がある。こうした目的のために、まず、対象企業の絞り込みを行う。具体的には、分析対象期間である2006年5月から2014年4月までの間に、予測精度の検証においてアウトカムとする不正イベントのネガティブ（不正無し）レコードが少なくとも5レコードは存在する企業のみを残す。この結果、当該期間においてネガティブレコードが4回以下の数しか記録されていない企業は共分散構造の推定において用いられないこととなる。こうした処理は、信頼できる推定値を得るためには不正を含まないレコードが時系列方向で一定数存在するもののみを用いるべきと考えるためである。次に、企業×科目毎の標準化に用いる平均値と標準偏差の計算を行う。具体的には、会社毎に分析対象期間である2006年5月から2014年4

月までの間におけるネガティブレコードのみを用いて、各企業×科目レベルでの平均値と標準偏差を計算する。最後に、これらの情報を用いて各企業×科目の数値を標準化されたスコアに変換する。このとき、標準スコアが定義できないもの（資本金など、数値に動きがないもの）については、標準スコアを0として記録する。

第五に、決算年月を基準に上記の処理を終えたデータを、モデル作成データとスコア計算データに分割する。具体的には、2006年5月から2014年3月までのデータをモデル作成データ（Model）とし、2014年4月から2018年10月までのデータをスコア計算データ（Score）とする。更に、「モデル作成データ」は企業単位で Train データと Test データへランダムサンプリングによる100種類の分割を行う。具体的には、モデル作成データのうち、1回もポジティブレコードのない企業を Train データと Test データについて各々80%、20%の割合で分割し、モデル作成データのうち、1回以上ポジティブレコードの有る企業は全て Test データとして用いる。つまり、訂正が一切存在しない healthy データのみで共分散構造の推定を行うことになる、ポジティブレコードのない企業については Train データと Test データの分割を100回行う。表3は以上のサンプリング方法について再度まとめたものである。

表3：サンプリング方法のまとめ

決算年月	Model/Score 分割	ポジティブ有無	Train/Test/Score 分割
'06/05 ~ '14/03	Model	ポジティブなし企業	Train : Test = 80 : 20
		ポジティブあり企業	すべて Test
'14/04 ~ '18/10	Score	すべての企業	すべて Score

こうした処理の結果、表4-1記載のモデル作成データと表4-2記載のスコア対

象データが構築される。

表 4 - 1 : モデル作成データ ('06/05 ~ '14/03) の概要

不正モデル	総企業数	ポジティブ レコード なし企業数	ポジティブ レコード あり企業数	レコード総数	ポジティブ レコード数
売上	1,761	1,754	7	12,702	13
費用	1,761	1,756	5	12,697	8
資産	1,761	1,752	9	12,714	25

表 4 - 2 : スコア対象データ ('14/04 ~ '18/10) の概要

不正モデル	総企業数	レコード総数	ポジティブレコード数
売上	1,529	5,674	8
費用	1,529	5,761	5
資産	1,529	5,763	7

構築されたこれらのデータのうち Train データを用いて、第一に、graphical lasso を用いた共分散構造の推定を行う。具体的には、100 回のサンプリング毎に抽出された Train データをモデル作成に使用する。この時、各サンプリングに対して、正則化項に付随するハイパーパラメータに関して 27 通りのセットアップを準備する。本稿では、試行的な実験結果を踏まえて、 $[1 - 9] \times 0.001$ 、 $[1 - 9] \times 0.010$ 、 $[1 - 9] \times 0.100$  の 27 通りの値を用いる。当該パラメータが大きくなるにつれてペナルティが大きくなるため、より強いスパース化が行われる。各サンプルとハイパーパラメータを所与として、科目間の精度行列  $\Lambda$  が推定される。

第二に、この推定行列と Test データを用いて各科目の異常スコアを計算する。具体的には、100 回のサンプリング毎とペナルティのペア毎に得られた精度行列を所与として、モデル作成時に参照した科目全ての Test データにおける異常スコアを以下の定義に基づいて計算する。

$$a_i(\mathbf{x}) = -\frac{1}{2}\Lambda_{i,i} \left\{ x_i - \left( m_i - \frac{1}{\Lambda_{i,i}} \sum_{j \neq i}^M \Lambda_{i,j} (x_j - m_j) \right) \right\}^2 - \frac{1}{2} \ln \frac{\Lambda_{i,i}}{2\pi}$$

ここで、 $M$ は科目数、 $x_j$ は異常スコアの計算対象となる任意のレコードに関する科目  $i$ の標準化済みスコア ( $1 \leq i \leq M$ )、 $m_j$ は科目  $i$ の標準化済みスコアのサンプリングごとの平均値であり現在のセットアップでは0となる。 $\Lambda_{i,j}$ は既述の精度行列である。

第三に、不正検出に最適なペナルティの大きさを科目ごとに決定する。具体的には、Test データを用いて、全ペナルティの異常スコアと不正フラグ（売上過大/費用過小/資産過大）の AUC を科目ごとに計算したうえで、科目毎に AUC が最大となるペナルティを採用する。また、代替的な方法として、科目ごとではなく、全科目の異常スコアの平均値/最大値を各レコードの代表値とする方法も用いる。

最後に、以上のプロセスで同定された最適ペナルティと対応する精度行列、Score データを用いて、各科目の異常スコアを計算する。

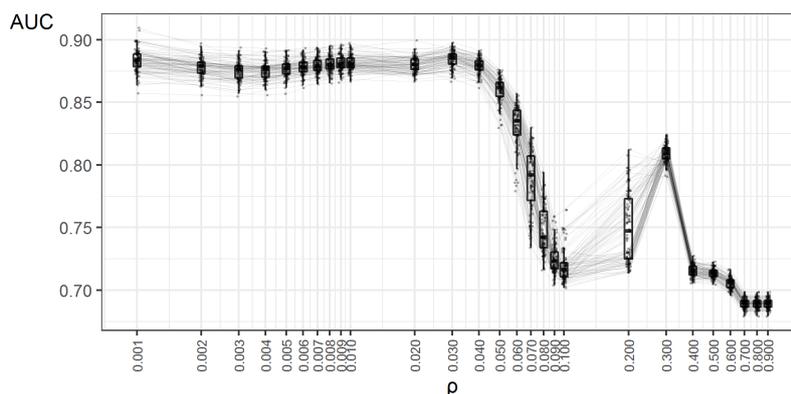
#### 4. 実証結果

前節で整理した手順に従い、Train データを用いてモデルを作成し、Test データを用いた全科目の異常スコア計算と科目毎の不正フラグ予測に関する AUC 計算からなるプロセスを 100 回のランダムサンプリングの個々の試行に関して行い、AUC が最

大となるペナルティ ( $\rho$ ) を決定したうえで、当該のペナルティ ( $\rho$ ) と Score データから計算した異常スコアについて、不正フラグ予測に関する AUC を再度計算した結果は以下の通りである。具体的には、不正を検知する対象として設定した科目のうち売上不正と資産不正に関する結果を示す。

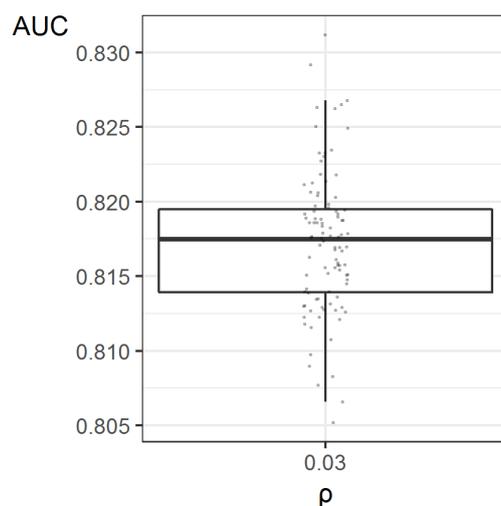
第一に、売上不正の検出を目的として売上総利益に着目した結果である。Test データに基づく推定から、売上総利益の異常スコアと売上不正フラグの AUC が最大となるペナルティ ( $\rho$ ) は 0.030 と見積もられた。図 1 は各  $\rho$  に対する AUC の挙動を図示したものである。

図 1 : Test データの AUC (売上総利益×売上不正)



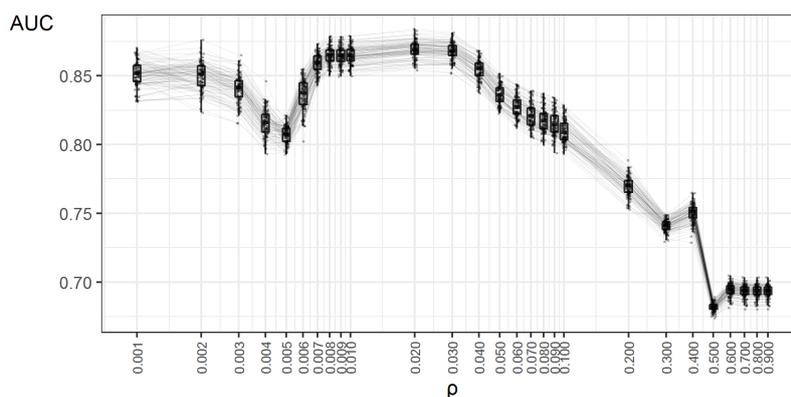
この値を所与として Score データから売上総利益の異常スコアを算出し、売上不正フラグに関する AUC を計測した。図 2 は 100 回の試行ごとに算出されたこれらの AUC を描画したものであり、概ね 0.815 ~ 0.820 程度の精度を実現していることが分かる。

図 2 : Penalty ( $\rho$ ) = 0.030 のとき、Score データの AUC (売上総利益×売上不正)



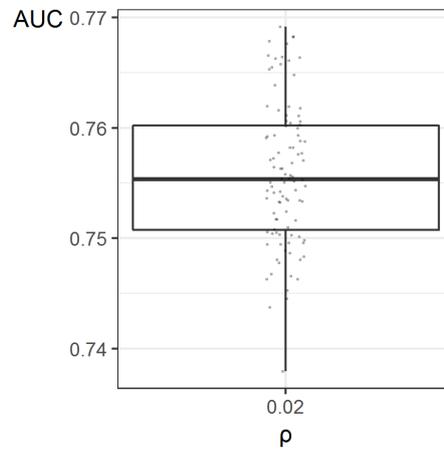
第二に、資産不正の検出を目的として純資産に着目した結果である。Test データに基づく推定から、純資産の異常スコアと資産不正フラグの AUC が最大となるペナルティ ( $\rho$ ) は 0.020 と見積もられた。図 3 は各  $\rho$  に対する AUC の挙動を図示したものである。

図 3 : Test データの AUC (純資産×資産不正)



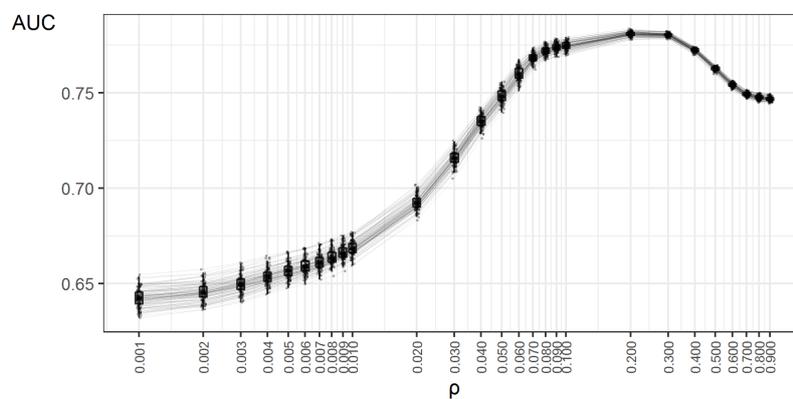
この値を所与として Score データから売上総利益の異常スコアを算出し、資産不正フラグに関する AUC を計測した。図 4 は 100 回の試行ごとに算出されたこれらの AUC を描画したものであり、概ね 0.750~0.760 程度の精度を実現していることが分かる。

図 4 : Penalty ( $\rho$ ) =0.020 のとき、Score データの AUC (純資産×資産不正)



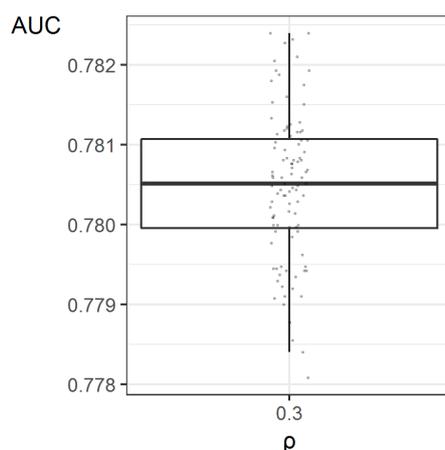
第三に、資産不正の検出を目的とした分析において、純資産単体の異常スコアに着目するのではなく、全科目の異常スコアの平均値に着目した結果である。Test データに基づく推定から、全科目の異常スコア平均値と資産不正フラグの AUC が最大となるペナルティ ( $\rho$ ) は 0.300 と見積もられた。図 5 は各  $\rho$  に対する AUC の挙動を図示したものである。

図 5 : Test データの AUC (全科目の平均×資産不正)



この値を所与として Score データから全科目の平均異常スコアを算出し、資産不正フラグに関する AUC を計測した。図 6 は 100 回の試行ごとに算出されたこれらの AUC を描画したものであり、概ね 0.780 程度の精度を実現していることが分かる。

図 6 : Penalty ( $\rho$ ) = 0.300 のとき、Score データの AUC (全科目の平均×資産不正)



以上の結果は、Model データ内で推定された共分散構造に関するパラメータと同データ内における最適化の結果として同定されたハイパーパラメータから構成される異常度計測モデルが、out-of-sample データである Score データにおいて、不正の検知に関して実務的に見ても十分な水準の精度 (0.75 から 0.8 程度の AUC) を実現していることを意味する。本研究で構築したモデルは、Model データ内の healthy データを用いて各科目間の共分散構造を推定しているという意味で **unsupervised machine learning** としての位置づけが可能であるが、重要なハイパーパラメータである正則化項の設定においては Model データ内のポジティブデータも併せて参照しているという意味では **supervised machine learning** としての整理も可能である。ハイパーパラメータの設定に当たって、このような形で実運用時における関心の対象となる予測対象に紐づける形で最適化を行うことの重要性は先行研究でも強調されているところであるが (例：井

手 2017)、本研究での実証結果はこうしたある種のハイブリッドなモデル構築が実務的に有用であることを示唆している。

## 5. ディスカッション

本節では、本稿での実証結果を踏まえたディスカッションを行う。第一に、本モデルで作成した共分散構造は、モデル作成対象期間における変数系列間の関係性を示すものであるが、モデル作成で特定された関係性が常に維持されるとは限らない。例えば、経済危機や自然災害といった要因によって変数系列間の関係性が変化し、コンセプトドリフトが生じうる。この結果、算出した異常スコアの信頼性が低下しうる。コンセプトドリフトへの対応策として、**graphical lasso** モデルが陳腐化していないか継続的にモニタリングすることが重要である。

第二に、**graphical lasso** モデルは、各変数系列が従う多変量正規分布の共分散構造を求めているが、本検討で使用した財務データが常に正規分布に従うとは想定し難い。本検討の目的が共分散構造の特定ではなく異常検知であるという見地に基くと、財務データの分布として正規分布を仮定しても異常検知を行う目的は実現される可能性がある。しかし、財務データへ **Box-Cox** 変換等を施した上でより信頼性が高いと考えられるモデル構築を行い、異常検知の精度にどのような変化が生じるかを確認することが今後の課題として考えられる。

第三に、不正を行う企業に着目すると、実務上、単一企業のみで不正が完結することは少なく、循環取引のように複数企業が結託して行う場合もあり得る。このような不正を検知・予測するためには、勘定科目レベルや企業レベルと言った粒度ではなく、取引関係を考慮に入れたより大きな粒度での検知・予測が必要となり得る。

第四に、様々な粒度で高い精度のスコアが提供されることを前提にすると、機械に

よる判断と人による判断とをどのように組合せ、より適切な判断へと導くべきかが論点となり得る。より効果的・効率的な事業運営のためには、機械と人間の向き不向きを踏まえた業務の再構築が必要となり得るであろう(例:Miyakawa and Shintani 2020)。

## 6. まとめ

本研究では、本邦上場企業の不正会計イベントを対象として、勘定科目レベルでの異常検知(不正検知)を試みた。具体的には、変数間の偏相関関係の推定を、次元圧縮を行いつつ実施するスパースモデリング手法である **graphical lasso** を用いて科目間の共分散構造を推定したうえで、各科目の条件付き分布を前提とした異常度計測を試みることによって、勘定科目レベルの不正会計検知手法を実証的に検討した。構築した異常度スコアの **out of sample** での検知能力をテストしたところ、実務的な観点から十分な検知精度を達成したほか、スパース化による検知精度の改善も図られていることが分かった。

参考文献

井手剛 (2017) 「依存関係にスパース性を入れる」 岩波データサイエンス vol.5: 48-63.

岩波書店

井手剛, 杉山将 (2015) 異常検知と変化検知. 講談社サイエンティフィク.

宇宿哲平, 近藤聡, 白木研吾, 菅美希, 宮川大介. 2019. 「機械学習手法を用いた不正会計の検知と予測」 RIETI Discussion Paper Series 19-J-039.

宇宿哲平, 近藤聡, 白木研吾, 宮川大介, 柳岡優希. 2021. 「機械学習手法を用いた不正会計予測: 非上場企業データを用いた検討」 Hitotsubashi University FS Working Paper Series 2021-J-001.

Banerjee, O., L. E. Ghaoui, A. d'Aspremont, and G. Natsoulis. 2006. Convex Optimization Techniques for Fitting Sparse Gaussian Graphical Models. ICML '06: Proceedings of the 23rd international conference on Machine learning: 89-96.

Dechow, P. M., Ge, W., Larson, C. R. and R.G. Sloan. 2011. Predicting Material Accounting Misstatements. *Contemporary Accounting Research* 28(1): 17-82.

Dechow, P. M., Ge, W., and C. Schrand. 2010. Understanding Earnings Quality: A Review of the Proxies, their Determinants and their Consequences. *Journal of Accounting and Economics* 50 (2)(3): 344-401.

Dechow, P. M., R. G. Sloan, and A. P. Sweeney. 1996. Causes and Consequences of Earnings Misstatement: An Analysis of Firms Subject to Enforcement Actions by the SEC. *Contemporary Accounting Research* 13(1): 1-36.

Friedman, J., T. Hastie, and R. Tibshirani. 2008. Sparse Inverse Covariance Estimation with the Graphical Lasso. *Biostatistics* 9(3): 432-41.

Idé, T., A. C. Lozano, N. Abe, and Y. Liu. 2009. Proximity-Based Anomaly Detection using Sparse Structure Learning. Proceedings of the 2009 SIAM International Conference on

Data Mining (SDM): 97-108.

Miyakawa, D. and K. Shintani (2020), “Disagreement between Human and Machine Predictions”, IMES Discussion Paper Series 2020-E-11, Bank of Japan.

Song, M., N. Oshiro, and A. Shuto. 2016. Predicting Accounting Fraud: Evidence from Japan. *The Japanese Accounting Review* 6: pp.17-63.