

# データ科学と天文学

池田 思朗<sup>1</sup>  
森井 幹雄<sup>2</sup>

〈統計数理研究所〒190-8562 東京都立川市緑町 10-3〉  
e-mail: <sup>1</sup> shiro@ism.ac.jp, <sup>2</sup> morii@ism.ac.jp



池田



森井

最近、機械学習や人工知能といったデータ科学分野の進展が目覚ましい。この20年ほど、応用数学分野の理論的進展と計算機の処理能力の飛躍的な向上が相まって、データ解析の手法が発展した結果である。こうした新たなデータ処理の方法は産業を通じた社会への貢献だけでなく、自然科学の分野にも影響を及ぼしていくはずである。実際に生物学をはじめ、すでに大きな変化をした分野が存在する。本稿では、天文学において人工知能や機械学習、統計学などのデータ科学分野の手法が今後どのような役割を担っていくのか考えていく。

## 1. はじめに

「データ科学と天文学」に関する記事が天文月報にシリーズとして連載されることになった。本稿では、その第一回として天文学とデータ科学との関わりについていくつかの例を交えながら簡単に紹介したい。

データ科学はデータを対象とする科学であり、目的の情報をデータから効率的に取り出すための方法を研究の対象とする。これは、以前から統計学や信号処理学などが扱っていた問題である。そこに計算機科学や最適化理論の方法が取り入れられ、機械学習や人工知能といった新しい分野と渾然一体となって新たな名前と呼ばれるようになってきた。GoogleをはじめとするIT大企業はこうしたデータ科学の担い手であり、その恩恵を受けているといえる。国内でもこの数年、データ科学に関する書籍が数多く出版され、AIベンチャー企業や国が主導する人工知能の研究所が複数できている。

この新しい科学の影響は他の学問分野にも及んでいる。生物学ではゲノム配列が解読できるよう

になって、すなわちデータがデジタル化されてから急速にデータ科学の方法が使われるようになり、バイオインフォマティクスやシステムバイオロジーと呼ばれる分野が生まれた。脳科学や材料科学の分野でもデータ科学の方法が導入され、革新が起きようとしている。

天文学はというと、あらゆる波長の観測結果がデジタルデータとして保存されるようになったのはずいぶん前のことだ。それ以来、計算機上で統計的な解析が行われてきた。今の状況を見ると、多くの解析では20年、あるいはそれ以前に確立された手法を修正しながら使い続けている、という印象を受ける。最新の方法を用いた個別の研究は散見されるものの、分野全体としてデータ科学の影響を大きく受けたというにはまだ早いだろう。

現状はともかく、LSST (Large Synoptic Survey Telescope) やSKA (Square Kilometre Array) といった例を見れば明らかのように、10年先を考えると新たなデータ解析手法、つまりデータ科学の導入は必要不可欠である。このシリーズを通じて新たな流れを感じていただきたい。

## 2. 天文学がデータ科学に期待すること

これまで、いくつかの研究を通じて天文学者と議論を重ねてきた。その内容をまとめると、天文学からデータ科学への期待、言い換えるならばデータ科学から天文学への貢献の可能性には大きく二つの方向性があると感じている。

一つは、新たな方法による解析結果の改善である。限られたデータからデータ科学の方法によってより多くの情報が得られれば天文学への貢献は大きい。このシリーズでは植村氏の変動天体の解析や、秋山氏の電波干渉計のイメージング法、平野氏の系外惑星の検出はこうした方向性の例としていだろう。この十数年の間、データ科学ではいくつかの重要な要素技術が提案されているが、その中でもスパースモデリングの方法は重要なものである。植村氏、秋山氏の研究でもこうした方法が重要な役割を果たしている。本稿ではスパースモデリングの概略を示しておく。

もう一つの方向性はビッグデータへの対応である。今後、観測機器の高性能化によって高精度なデータが膨大に得られるようになると、天文学者がこれまで解析に用いてきた方法では対応しきれないこともありうる。その場合にはデータ解析の方法も更新する必要があるだろう。例えば、天体の発見などを目視に頼っていた場合、処理できるデータの量には限界がある。ここに新しいデータ科学の方法が求められている。本稿では、われわれが行ってきた研究の中から、Tome-e Gozen<sup>1),2)</sup>のデータ圧縮に関する方法<sup>3)</sup>と、すばる望遠鏡のHSC (Hyper Suprime-Cam) のデータから超新星を発見するために行った試み<sup>4)</sup>を示す。

## 3. スパースモデリング

スパース性を用いた情報処理は1990年代に提案された。この新しい話題はさまざまな理論的な問題を提供し、今世紀に入った頃には、統計学や

信号処理学にはじまり、情報理論、最適化理論などほとんどの応用数理の分野で関連した研究が進められた。スパースモデリングという言葉がいつから使われるようになったのかははっきりしないが、少なくとも国内では岡田真人が代表を務めた科研費新学術領域が始まったあとである。

天文学など、スパースモデリングを用いる立場からは、スパースモデリングの方法は主に不良設定問題、いわゆる逆問題を解く方法として認識されているだろう。簡単な例として次の問題を考える。

$$\begin{aligned} \mathbf{y} &= \mathbf{A}\mathbf{x}, \\ \mathbf{y} &\in \mathbb{R}^M, \quad \mathbf{x} \in \mathbb{R}^N, \quad \mathbf{A} \in \mathbb{R}^{M \times N} \end{aligned} \quad (1)$$

ベクトル $\mathbf{y}$ と行列 $\mathbf{A}$ が既知であるとき $\mathbf{x}$ を求める問題はなじみ深い連立一次方程式である。特に $M < N$ のときを考えると、一般に式(1)を満たす解は無限に存在し、一意に決めることができない。解は不定である。

解を一意に決めるには、 $\mathbf{x}$ に条件を課し、解の候補を制限すればよい。スパースモデリングでは「できるだけ0の多い」すなわちスパース (sparse) な解を探すことにする。特に、上記の連立一次方程式にスパース性を導入した定式化はCompressed Sensing (圧縮センシング) と呼ばれている<sup>5)</sup>。圧縮センシングでは真の $\mathbf{x}$ がスパースであり、観測にどのような条件があれば解が求まるのか、そして具体的にどうやって求めるか、などを議論する。

注意してほしいのだが、スパースモデリングの方法は真の解のスパース性を確かめる方法ではない。「真の $\mathbf{x}$ がスパース」だということは仮定であって、対象がどのようなスパース性をもっているかはスパースモデリングを使う側が責任をもつ必要がある。

具体的な解の求め方についても少し触れておく。 $\mathbf{x}$ に0が多いことがわかっていても、それを求める方法は自明ではない。 $\mathbf{x}$ の成分のうち、せいぜい $K$ 個の成分が非零だとしても、どの

成分が非零かがわからなければ、 $N$ 個の成分から  $1 \sim K$  の成分を選び出す組み合わせを網羅的に調べ上げる必要があり、 $N$  が大きければ現実的な計算時間では解けないからだ。

圧縮センシングが成功した理由は、 $A$  がある条件を満たし、 $\mathbf{x}$  のスパース性が高ければ  $\mathbf{y} = A\mathbf{x}$  を満たす  $\mathbf{x}$  のうち  $\|\mathbf{x}\|_{\ell_1} = \sum_{i=1}^N |x_i|$  を最小とする解を求めることによって、正しいスパースな解が求まると示したことだ<sup>6)</sup>。最適化の問題として書かならば、以下の問題を解くことになる。

$$\min_{\mathbf{x}} \|\mathbf{x}\|_{\ell_1}, \text{ subject to } \mathbf{y} = A\mathbf{x}. \quad (2)$$

最適化理論では解き難い問題を解ける形に変えた問題を「緩和問題」と呼ぶ。上記の問題も「もっともスパースな解を求める」という問題の緩和問題である。緩和問題の解は一般には本当に求めたい解と一致しないが、圧縮センシングの場合には一致する場合があることが幸運だったといえる。 $\|\mathbf{x}\|_{\ell_1}$  の最適化は線形計画法という最適化問題となり、大規模な問題でも現実的な時間で解ける。

一般の測定では観測に雑音加わる。その場合、式(1)、あるいは(2)を用いるのは適切ではない。そこで、いわゆる  $\chi^2$  誤差である  $\|\mathbf{y} - A\mathbf{x}\|_{\ell_2}^2$  と  $\|\mathbf{x}\|_{\ell_1}$  とのバランスをとって解を求める方法が提案された。すなわち、以下のような関数を最適化問題の対象とする。

$$\|\mathbf{y} - A\mathbf{x}\|_{\ell_2}^2 + \lambda \|\mathbf{x}\|_{\ell_1} \quad (3)$$

$\lambda$  は二つの項のバランスをとる項である。この問題は Least Absolute Shrinkage and Selection Operator (LASSO) と呼ばれ、1996年に Tibshirani によって提案された<sup>7)</sup>。LASSOの問題は二次計画法と呼ばれる最適化問題となる。近接勾配法などを用いて、数万次元の問題であっても解を求めることが可能である。

統計学で線形回帰 (linear regression) と呼ぶモデルでは、観測変数  $\mathbf{y}$  と従属変数  $A$  との間に線形関係を仮定し、 $\mathbf{x}$  にあたる係数を求めること

を目標とする。線形回帰は統計学の基本的なモデルであって、古くから各分野で用いられている。LASSOは線形回帰が不良設定問題となる場合を解決する方法として提案され、ゲノムを用いた医学分野で広く用いられるようになった。天文学でも線形回帰分析はしばしば用いられており、LASSOによる解析を行った結果も報告されている<sup>8)</sup>。

線形回帰を仮定するまでもなく、 $\mathbf{y}$  と  $\mathbf{x}$  の関係が線形となる問題は多く存在する。天文学にもよく現れるフーリエ変換は良い例である。本シリーズの植村氏原稿では変光星のライトカーブの周期解析<sup>9)</sup>、秋山氏原稿では電波干渉計イメージング<sup>10)</sup> にスパースモデリングの方法を用いている。どちらもフーリエ変換によって、観測量と求める量との間に線形関係が定義されている。

式(2)や(3)で示された問題では、 $\mathbf{x}$  というベクトルに関するスパース性を利用した方法を定義した。スパースモデリングでは、こうしたベクトルの問題だけではなく、行列やテンソルの問題も提案され、多くの研究が行われている。行列のスパース性としては成分に0が多いというスパース性もちろん考えられるが、それとは別にもう一つ、行列のランクが小さい、言い換えれば、特異値 (正方行列ならば固有値) の次元が小さい、というスパース性も考えられる。こうした二つのスパース性を組み合わせ、行列に対するスパースモデリングの方法が提案されている<sup>11)</sup>。この方法を応用した例については次章で紹介する。テンソルへの拡張についてはここでは触れないが、機械学習での研究は盛んであることから、興味があれば文献を見て欲しい<sup>12)</sup>。

#### 4. Tomo-e Gozenにおける動画データの圧縮法

前章に述べたように、行列に対するスパースモデリングの応用として、Tomo-e Gozenで得られた動画データの圧縮手法について紹介する。

Tomo-e Gozenは、東京大学木曾観測所の口径1.05 m シュミット望遠鏡に設置されている超広視野カメラである。撮像素子として可視光天文学でよく用いられる CCDではなく CMOSを用いることにより、CCDでは難しい10秒程度以下の時間分解能で撮像することができる。このカメラは、84枚の2k×1kピクセルCMOSチップにより構成され、20平方度の視野をカバーしながら2 Hzの高頻度で撮像する。これにより、サブ秒から数十秒のタイムスケールで変動する天体現象の観測や突発天体の発見を目指している。新たな天文学「動画天文学」の幕開けである。

Tomo-e Gozenは素晴らしい観測成果を上げると期待されるが、問題はそのデータ量である。一晩の観測で約30テラバイトもの動画データが取得できる。科学的成果を上げるには、この膨大なデータの中から効率的に変動天体や突発天体を検出しなければならない。また、継続的に運用していくためにはデータ圧縮も行うべきだ。われわれは、行列に対するスパースモデリングの方法を用いて突発現象を保持しながら約10分の1にデータを圧縮する方法を提案した。

さきほど述べたように、行列のスパース性として、ランクが低い低ランク性と成分に0が多いというスパース性を考えることができる。ここではこの二つのスパース性を使い、行列を低ランク行列とスパース行列の和に分解する方法を用いる。すなわち、スパース性を用いた行列分解である。Candèsらによって提案されたRobust PCAでは、行列の低ランク性は、行列の核ノルムを小さくすることによって、スパース性は $l_1$ ノルムを小さくすることで実現する<sup>11)</sup>。この関数は凸であるため最適化計算は容易である。同様の手法にGoDec<sup>13)</sup>という手法があり、計算が速いことと制約の付け方が便利であるため、Tomo-e Gozenのデータに対してはGoDecを用いた。

天体の動画データは、天球面の2次元と時間方向1次元からなる3次元データである。まず、動

画データの各フレーム $n_x \times n_y$  pixel<sup>2</sup> (2k×1k)のデータを縦に並べ、時間方向( $n_f=400$ フレーム)は横に並べるとサイズが $n_x n_y \times n_f$ の行列が作られる(図1)。定常天体は明るさと位置が一定なので、この行列中においてはほぼ同じ値が1行に並ぶことになる。一方、突発天体は行列中にスパースな要素として表現される。

この行列に対して行列分解を行うと、 $L+S+G$ に分解される。ここで、 $L, S, G$ はそれぞれ低ランク、スパース、ノイズ行列である。 $L$ はさらに特異値分解(SVD)によって $L=UDV^T$ に分解される。ここで、 $U$ と $V$ は直交行列、 $D$ は対角行列である。図2は、 $(U, D, V)$ を表示したものである。 $U$ の第1列目に相当する1番目のイメージに、 $V$ の1行目を掛け算すると、最も主要な成分からなる動画が構成できる。同様に2, 3, …番目の成分の動画が構成できる。それぞれの動画のウエイトは、対角行列 $D$ の対角成分、すなわち特異値の大きさと表すことができる。各成分の動画に特異値のウエイトを掛けて足したものが低ランク動画 $L$ である。

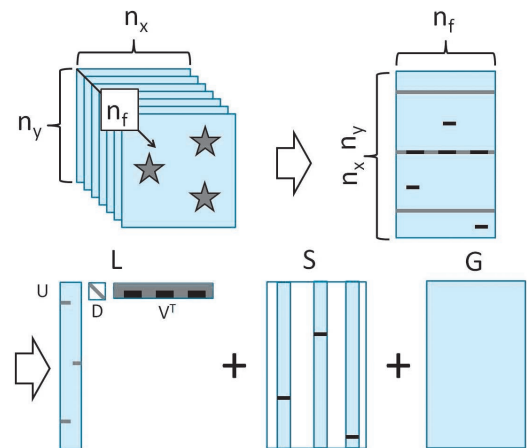


図1 低ランク行列分解を用いたデータ圧縮: 観測動画は行列に変換されGoDecによって $L+S+G$ に分解される。 $L$ はさらにSVDによって $UDV^T$ に分解される。この図は定常天体が二つ、変動天体が一つと突発天体が三つ含まれる例を示している(文献3の図1を改変)。



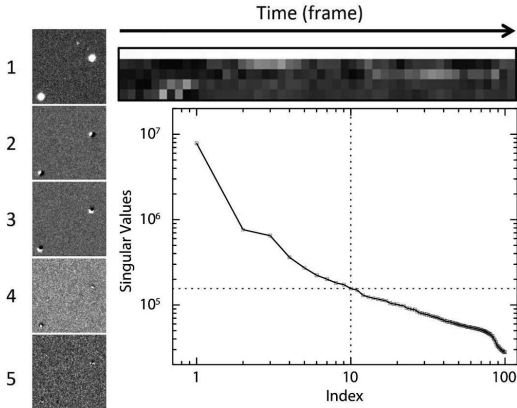


図2 低ランク動画の構造: 左側は,  $U$  の最初の5列に相当するイメージ. 上のパネルは, 対応する  $V$  の5行. 真ん中のグラフは特異値の大きさ (文献3より抜粋).

図2の左にあるイメージを見ると明らかなように, 5番目の動画以降はほぼノイズだけになる. 例えば特異値の値を10番目まで残すことにし, 残りの10から400番目の特異値に対応する動画は消去することによって大幅にデータが圧縮できることになる.  $S$  については, 突発天体が発生する頻度は多くないため, 何らかの突発現象が検出されるのは, 400フレーム中でせいぜい数枚であろうと予想される. 突発天体が検出されたフレーム以外は消去してしまって構わない. このような処理により, 突発現象をスパース行列に保持しつつ, 約10分の1にデータが圧縮できる.

Tomo-e Gozenのプロトタイプ機で取得した動画データに対して, 実際にこの処理を施したイメージを図3に示す. 上のパネルには実際に観測された短時間だけ光った点源が映っている. このような天体の正体については言及しないが, Tomo-e Gozenの潜在能力を示す一例である. 下のパネルには流星が映っている. スパース動画のイメージの中にこれらの突発現象が抽出できているのに対して, 低ランク動画のイメージからは消えている.

従来の手法であれば, まず天体を Point Spread

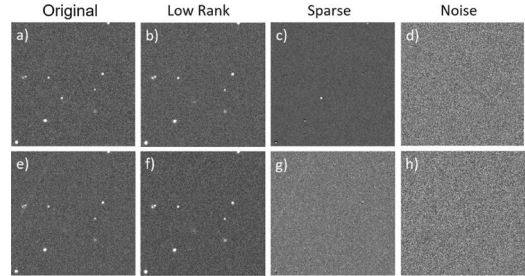


図3 Tomo-e Gozenプロトタイプ機により取得された動画に対して, 低ランク行列分解を行った結果: 上側 (c) には突発点源, 下側 (g) には流星 (上から左の方向に直線状に伸びている) が映っている (文献3より抜粋).

Function (PSF) を元に検出・測光し, ライトカーブを描くことにより, 突発天体を検出する. しかし, それでは多くの手順が必要となり, 計算に時間がかかるし, プログラムを作り込むために大きな労力を割くことになる. また, 流星のような点源ではない突発現象を検出するには, そのための処理が別途必要になる. それに対して, 低ランク行列分解手法では, 天体形状の情報を全く使わなくても突発現象が抽出できる点が面白い.

汎用のPC (Intel Xeon 3.70 GHz, 8コア, メモリ60 GB) を用いて, Tomo-e GozenのCMOSチップ1枚 (1k×2k) の200秒 (400枚分) のデータを処理したところ, 実観測時間の1.6倍かかることを確認した. 観測は夜だけなのに対し, データ処理は一日中可能なので一晩のデータを一日で処理できる. プログラミングを工夫すれば, 数年のうちに運用も可能だろう.

## 5. HSCデータからの超新星の選別

われわれは, すばる望遠鏡の超広視野主焦点カメラ (HSC: Hyper Suprime-Cam) を用いたサーベイ観測「すばる戦略枠プログラム」のデータ解析に参加している. このプログラムは2014年に開始され, 5年間で300晩の観測を行う. 最終的に観測データはペタバイトに迫る分量になると見積もられている. 遠方銀河の分布を測定したり,

同じ領域を何度も観測することにより遠方で発生するIa型超新星を大量に検出する。これにより、ダークマターやダークエネルギーといった現代物理学最大の謎を解明する手掛かりを得るのが目的だ。そのためには、Ia型超新星を検出だけでなく発見後直ちに分光観測を行って、天体の赤方偏移量を測定しなければならない。分光観測は超新星が明るいうちに行う必要があり、人間が判別を行う時間的な余裕はあまりない。

ここでは、人間の代わりに機械学習を用いて変動天体を自動的に検出するシステムを構築した結果を紹介する<sup>4)</sup>。超新星のような変動天体を検出するには、過去に取得した参照画像と新たに取得した画像との間で差分をとれば良い。しかし実際には、図4のような偽物（ゴミ）が大量に残ってしまう。このようなゴミの数は、期待されている変動天体数の数百倍にも及ぶ。これらは原因がわかっているのだから、一つひとつ丁寧にプログラムを作り込めば理屈の上では除去可能である。しかしながら、これらに対応するためには相当な労力を割かなければならない。このような問題に対しては、機械学習が有効である。

HSCを用いた超新星探索の先行研究では、差分イメージをPSFでフィットしパラメータを10個ほど抽出し、それらの情報とイメージを元に人間が本物・偽物の判定を行っていた。そこで、そのパラメータ（特徴量）を入力とし本物・偽物の判定結果を返す関数（マシン）を機械学習を用いて作成することにした。われわれは、良い性能を示すことが知られている三つの機械学習手法（ブースティング、ランダムフォレスト、深層学習）を用いた。図5はそれぞれの手法で得られた判別性能（ROC（Receiver Operating Characteristic; 受信者動作特性）曲線）を表している。これらの方法はどれも同程度の性能になっていることがわかる。慣れていない読者のために図6にROC曲線の意味を説明したのでご覧いただきたい。

さて、三つの判別器にも得手不得手があり、選

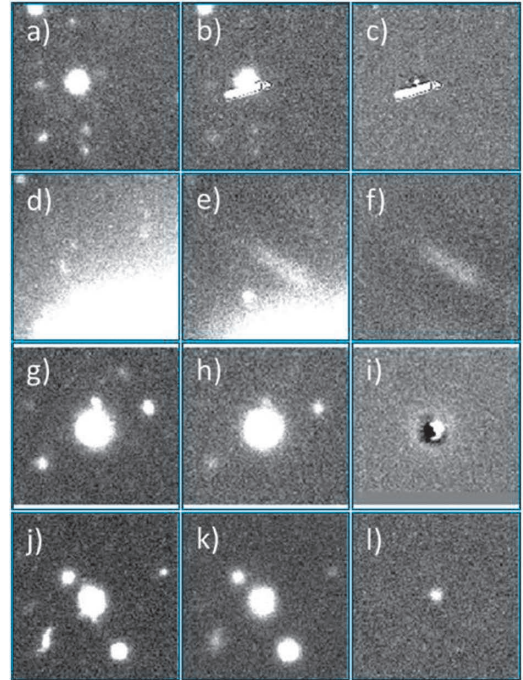


図4 Subaru/HSCの解析パイプラインの処理後に残った突発天体候補: 左, 中, 右側はそれぞれ, 参照画像, 新画像, 差分画像. c), f), i) はそれぞれ, 宇宙線, 明るい天体の周辺に現れるゴースト, PSFの違いによって生じた残差である. l) は本物の突発天体 (文献4より抜粋).

び出される天体の傾向は異なると考えられるため、これらの判別器の結果を統合する必要がある。そこで、三つの判別器の結果に対して3通りの組み合わせで統合した結果を求めることにした。三つを均等に扱うとしたとき、それらの組み合わせには、「多数決」、「全会一致」、「少数意見の尊重」の3通りがありうる。それらについてROC曲線を描いたところ、多数決が最も良い性能を示すことがわかった。また、現実の観測では、学習データと全く同じ環境でデータが取得されることはないため、必ず性能は悪くなる。異なる環境であってもそこそこ良い性能を示すRobust性が求められるが、われわれの判別器はある程度Robustであることも確かめられた。これらの判別器は、すでに、すばる望遠鏡の解析パイプラインにインス

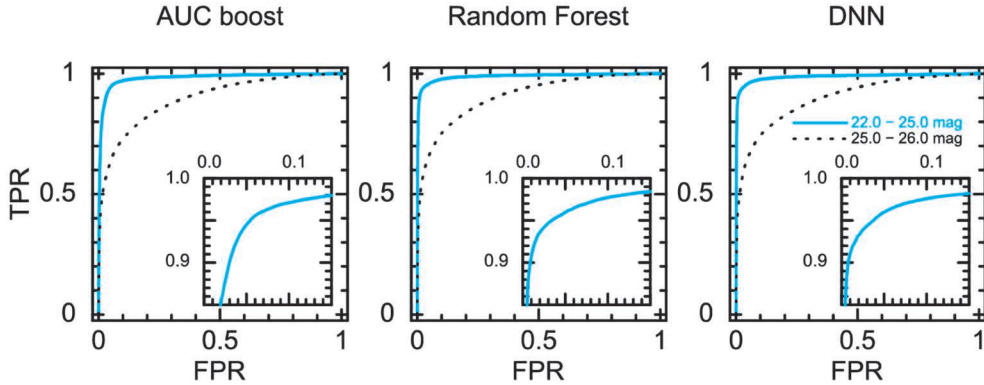


図5 超新星の判別問題に対する機械学習マシンの性能 (ROCカーブ): 右から順に, AUCブースティング, ランダムフォレスト, 深層学習の結果. ROC曲線の説明を図6に示す (文献4より抜粋).

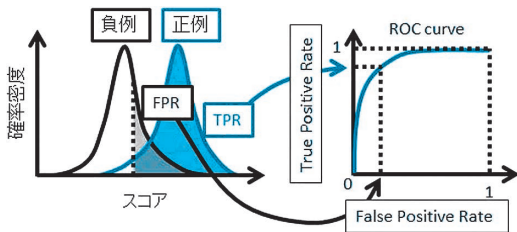


図6 判別問題の性能はROC曲線で評価できる. 判別器は特徴量ベクトルを入力としてスコアを出力する. 高いスコアであるほど本物である可能性が高いので, 閾値よりも高いスコアをもつサンプルを本物とみなす. 閾値ごとに, True Positive Rate (TPR) と False Positive Rate (FPR) が求められるが, 閾値を変化させてそれらをプロットしていくとROC曲線が得られる. ROC曲線は左上にあるほど, 判別器の性能が良い.

ツールされ超新星の選別に活用されている.

このような機械学習の応用は, LSSTを考えれば明らかのように, 高性能化した天体観測機器を用いるようになった観測天文学の多くの分野で必要不可欠な技術となるであろう.

## 6. まとめ

天文学は, 人類の歴史の初期から発展してきた学問の一つである. 常に最先端の計測技術が投入され, 宇宙の理解が深められてきた. データの解

析方法に関して最先端の方法が使われるのも当然である. 本稿では, 天文学とデータ科学の関わり方という観点からいくつかの例を示した.

本稿では, 天文学からデータ科学に期待すること, という立場からデータ解析の高性能化やビッグデータへの対応に関する成果を示した. しかし, データ科学を研究する立場からの目標は, われわれが主導して天文学へ新たな提案をすることである. それができるとき, 天文データ科学と呼ばれる分野が確立するのだろう. 例えば, 本シリーズで西道氏が示すシミュレーションに基づく宇宙論は, そうした提案を行える可能性がある分野の一つなのだと考えている.

さて, 本稿で示した研究例をみて, 目の前の問題に適用してみたいという人もいるだろう. 最新のデータ科学の方法には, すでにパッケージとしてダウンロードできる形で提供されているものも多くある. そうしたものを自分の問題に試してみることも可能だ. うまくいけばよいが, 必ずしも簡単ではない. そのような方のために, 最後に意見を述べたい.

データ科学の難しいところは, 多くの場合, 目的を達成するための方法が複数あるように見えることだ. 似たような手法がたくさんあって, どれを選べばいいのかわからないかもしれない. われ

われは、データのもつ質と量、背景にある物理、そしてどのような情報を引き出したいか、といったいくつかの要素の組み合わせのうえで、適切な方法を決める。読者もそのような場合にはデータ科学の研究者と相談して研究を進めるとよいだろう。良いデータ科学者であれば、すべてを総合的に判断して適切なアドバイスを与えてくれるに違いない。

われわれもこれまで波長、粒子、観測装置を問わず、さまざまなデータ解析の相談を受けてきた。可能な限り話を聞き、どのように貢献できるかを考えている。今後10年の間に起きる天文学の変化の力になれば、このうえない喜びである。

## 謝 辞

本研究の一部は科学研究費補助金(25120008, 17K05395) および JST CREST, (JPMJcR1414) によるサポートを受けて行われました。

## 参考文献

- 1) Sako, S., et al., 2016, SPIE, 9908, 99083P
- 2) Ohsawa, R., et al., 2016, SPIE, 9913, 991339
- 3) Morii, M., Ikeda, S., Sako, S., Ohsawa, R., 2017, ApJ, 835, 1
- 4) Morii, M., et al., 2016, PASJ, 68, 104
- 5) Donoho, D. L., 2006, IEEE trans. Inform. Theory, 52 (4), 1289
- 6) Candès, E. J., & Tao, T., 2005 IEEE trans. Inform. Theory, 51(12), 4203
- 7) Tibshirani, R., 1996, JRSS B, 58(1), 267
- 8) Uemura, M., et al., 2015, PASJ, 67, 55
- 9) 植村誠, 加藤太一, 2018, 天文月報, 111掲載予定
- 10) 秋山和徳, 本間希樹, 池田思朗, 2018, 天文月報, 111掲載予定
- 11) Candès, E. J., et al., 2011, J. ACM, 58(3), 11
- 12) 富岡亮太, 2015, スパース性に基づく機械学習(講談社)
- 13) Zhou, T., & Tao, D., 2011, Intl. Conf. on Machine Learning

## Data Science and Astronomy

Shiro IKEDA and Mikio MORII

*The Institute of Statistical Mathematics, 10-3 Midori-machi, Tachikawa, Tokyo 190-8562, Japan*

Abstract: As the result of theoretical developments of applied mathematics and technical progress of computing technologies in the last 20 years, "Data science" emerged. This new field includes machine learning and artificial intelligence. This new science has made a great impact not only on the society but also on natural science. Biology is a good example. Here, we discuss the role of the data science in the future of Astronomy.