６．AIの数理と数学教育
ーAIの時代をよりよく生きるためにー

藤原毅夫（東京大学名誉教授）注1）

１．はじめに：AIをめぐる誤解

現在では、多くの人が何らかの形でAI（Artificial Intelligence、人工知能）を利用した経験があるだろう。あるいは、気付かないうちに利用しているかもしれない。例えば、筆者はこの原稿をWordで書いているが、その途中で文章や漢字のチェックが入っている。オンライン会議システムZoomは、多言語の要約機能を持っている。これらはすべてAIの機能である。

AIの能力が格段に上がったと評判になり、OpenAIのChatGPTが無料で利用できるようになったとき、筆者も試みに使ってみた。「統計手法『箱ひげ図』を開発したのは誰か？」と尋ねたところ、答えは見事に外れた。学習データに当該文献（あるいは信頼できる情報）が含まれていなかったためである。Wikipediaには、開発者名として誤った名前（しかも著名な数理科学者の名）が掲載されていた。注2）

その後１年ほど経ったとき、たまたま筆者自身の統計の本文献1）の中の一節を入力し、「源氏物語風に書き換えて下さい」と頼んでみた。すると予想以上に出来栄えが良かったため、枕草子風、奥の細道風、徒然草風など、いくつかの文体でも試してみた。いずれも、筆者が高校時代に学んだ古典文の特徴をうまく織り込みながら書き換えてくれた。つい調子に乗って、俳句を、文章の趣旨を生かしながら四季の趣を織り込んで作るよう指示したり、さらに漢文や漢詩にまで変換させて遊んだ。

俳句までは筆者にもそれなりに鑑賞でき、結果を判断できる。しかし、漢詩は本当に漢詩になっているのだろうか、単に漢字が一行に5つ並んでいるだけで良しとしたのではなかろうかという疑問を持った。

（１）AIは「考えている」のか？

私たちは、自身が分かっていることの正否は判断できるが、分からないことの正否は判断できない。これは当たり前だが極めて重要なことである。知らないことに関しては、相手の肩書（権威）や結果のもっともらしさを、正否と置き換えていることが多い。

AIの利用に関しても、我々は結論をAIに委ねがちである。そのAIを作っている人や企業が信用できるから、結果がもっともらしいから、等々の理由によるのであり、内容が正しいかどうかを十分には検討しないことが多い。我々が正しいと受け入れているのにはそれなりの理由、あえて云えば次のような誤解があるからではないだろうか。

誤解①：AIは　“理解している”
誤解②：AIの出力は　“正しい”
誤解③：AIは　“自分で学ぶ”

（２）今なぜこの問題を取り上げたのか

AIは政策や行政による重要な判断に関与するようになった。その結果、広い利用者が注意をする必要が出てきた。安易な判断は、責任の所在を不明確にし、公平性を担保できないからである。AIを使うためには、それが何をやっているのか、どのように問題を扱っているのか、問題によって得手不得手はあるのか、等を知っていることが重要である。

本稿では、AIの基本はいくつかの数学であるということを説明し、利用者には判断する力が必要であるという当たり前なことを基に、数学教育について考える。第２節では、AIの基本技術は数学が支え、AIの本体であるコンピュータの電子技術の進歩が欠かせないことを説明する。第３節では、AI時代の数学（数理）教育はこれまでと同じでいいのかを考える。第４節はまとめである。

２．AIを支える4つの数学と工学

（１）線形代数（表現の数学）

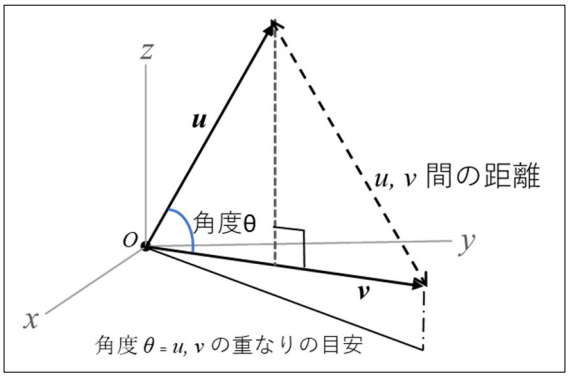
文書、統計データ、あるいは画像など、これらすべてを「情報」という。AIに情報が入力されると、それをデータとして取り扱うために、まず数値化される。

文書を考えてみよう。文書の文字は初めから

「コード化（数値化）」されている。この文字の並びを「語の集まり（分かち書き）」として認識していく。分かち書きにはいくつかの可能性があるので、最も自然な区切り（配列）を決定する。たとえば「ここではきものをぬいでください。」という文は、「ここで/はきもの（履物）を/ぬいで/ください/。」とするか、「ここでは/きもの(着物)を/ぬいで/ください/。」かを、前後の文脈から確率的に選び取る。

この作業では、「語」あるいは「文」は数値化されているわけである。抽象的あるいは一般的な名称を数値化し、いくつかの数値の組として表す。丁度、地図上の点の位置を、２つの数字「経度と緯度」で表すのと同じである。この2つの数字の組をベクトルという。「語」とか「文」の場合には、それらを評価する何千何万という数値が必要であるから、ベクトルの要素は何千何万となる。すなわち何千何万次元のベクトルで表される。

図１　２つのベクトル*u, v* の間の距離と重なり



ベクトルで表された特徴量を、「距離」と「重なり」で評価する。以下に簡単な計算例を示す（読み飛ばしてもよい）。

数値化の例を猫と子犬で見てみよう。それらの特徴を、例えば「生き物」軸（生物なら１、無生物なら０）、「大きさ」軸（巨大なら１、見えないほど小さければ０）、「可愛さ」軸（可愛ければ１、怖ければ０）、「懐き易さ」軸（すぐ懐くなら１，懐かなければ０）の上での値を用いて表す。この順番で猫と子犬を評価すれば、次のようになったとしよう。

[猫を表すベクトル]=（１，0.2，0.9，0.1），
[子犬を表すベクトル]=(１，0.2，0.9，0.9)．

ここでは、４番目の座標で初めて両者は区別されている。これはそれぞれを表す座標の場所の違いである。原点からの距離は

[猫]の距離=$\sqrt{1^2+0.2^2+0.9^2+0.1^2}$=1.36,
[子犬]の距離=$\sqrt{1^2+0.2^2+0.9^2+0.9^2}$=1.63

となる。さらに

[猫]と[子犬]の間の距離（違い）
=$\sqrt{(1-1)^2+(0.2-0.2)^2+(0.9-0.9)^2+(0.1-0.9)^2}$ = 0.8

となる。2点間の距離は「2つの概念の違い」の大きさを示す。それでは両者はどのくらい似ているだろうか。２つのベクトルの重なり具合は

猫・子犬=(1×1+0.2×0.2+0.9×0.9++0.1×0.9)/
(1.36×1.63)=0.88

と定義する。これは一般次元のベクトルの長さと重なり（内積）の定義によっている。「重なりはどのくらい似ているかという目安」であり、1に近いほど似ているといえる。ここでは4次元空間の中の点を扱ったが、実際には意味空間の次元は何千何万次元である。

このようにして、**対象の違いや類似性が数値として比較**できるようになる。ここで例示したのは、言葉の取り扱いである。統計データであれば、１つのデータはいくつかの特徴量で表現されているから、それはベクトルであり、まったく同じように処理される。

画像は２次元空間での点の集まりである。画像の色を３つ（RGB）に分解した上で、分解した画像の画素を、0，1の２値あるいはもっと細かい多値で表現することでデータが数値化される。そこに現れた図形の形や大きさを評価することにより、それらが何の画像であるかという判定がなされる。[注３）]

ベクトル空間という概念を用いて、3つの違った種類のデータ（文書、統計データ、画像）を、同じように取り扱うことができることを示した。このような大次元のベクトルの性質を扱う数学が**線形代数**である。従って、巨大なデータで表された統計を取り扱う場合にも、線形代数が必須であることが分かる。以上のような共通した取り扱いを可能にした線形代数を**「表現の数学」**と意味付けたのは、以上の理由による。

“AIが理解しているように見える”のは、それがこのような計算にのっとっているのであって、“私たちの理解”とは異質なものである。

（２）確率・統計（偶然を記述するための数学）

データをベクトルとして表現できたら、それを判断に結びつけなくてはならない。

統計を利用する立場では、観測事実からその原因となる事象が起こっている確率を推論する。

「たまたま（偶然に）」起きていることを、「確率」という言葉で記述する。しかし、確率というのは極めて不確実な土台に過ぎない。

従来の統計学（古典統計学）では事象の確率は既知で確定しているものと考えていたが、実際にはその確率自体が観測によって知られるものである。例を挙げてみよう。（ここは読み飛ばしてもよい。）

ある病気の罹患率は１万人に２人、すなわち何も検査情報がない段階で、この人が罹患している確率は0.02%である。検査で陽性と出たときに「本当に罹患している確率」はこれとは異なる。今病気の検査では罹患者に関して、90%の精度で実際に罹患しているかを判定できる。ただしこの検査では非罹患者に対して1.5%の確率で罹患者であるという判定がでる（擬陽性）。

この患者の最初の検査結果が、陽性と出たとする。この患者が陽性である確率は0.0002である。これを用いて本当に罹患しているかどうかを計算すると、本当に罹患している確率は0.0119（罹患率1.19%）、約１%であることが計算から分かる。すなわち、陽性と出たとしても、かなりの確率で本当の病気ではないことが分かる。検査の精度が90%と比較的高くても、検査結果の正確性はかなり低い値でしかない。ここでは**ベイズ推定**という統計的推定法を用いた。

検査で陽性と出ても実際に罹患している可能性は1.19%と低い。そこで、この患者に対して治療を開始する前に、もう一度同じ検査をする必要があると、担当医師は判断した。2回目の検査では、感染の確率を0.0119であるとしてベイズ推計を行うと、感染している確率は0.418、すなわち罹患率41.8%という結果が得られる。従って2回目の検査でも陽性という結果が得られた場合には、実際に罹患している確率は0.418、すなわち約42%まで上昇する。このように、情報を追加することで判断の確からしさが大きく向上されるのである。新しい情報によって確率を更新する方法を**ベイズ更新**という。

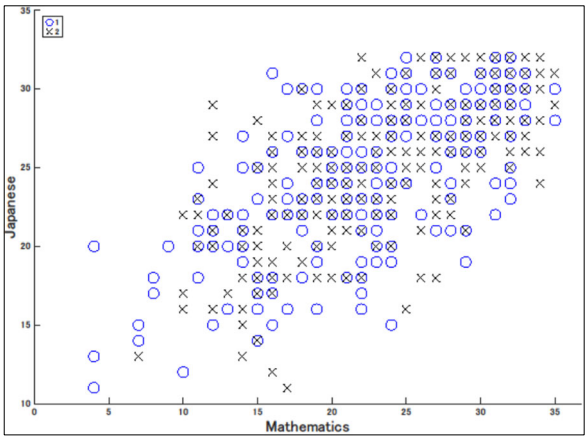
統計学のベイズ推定、ベイズ更新という手法を使ってサンプルの確度（ここでは罹患率）を上げていくことができる。[文献２）]

統計データは、（１）で述べたような数値データ（ベクトル）が常に一定値に収束するわけではない。例えば（１）の例の子犬に対するデータが必ず（１, 0.2, 0.9, 0.9）となるわけではない。ある子犬は、子犬のくせに大きかったり、あるいは丸まって寝てばかりで可愛くないということはいくらでもある。犬種によってはなかなか懐かない。だから、いろいろな子犬によって、個々のサンプルを示す点（値）は、どこかの範囲に分布している。分布がいくつかのグループに分かれていることもあるだろう。例えば小型犬種の場合と大型犬種の場合の様に。

ベイズ推定はデータ１点の確度を上げるために用いるばかりではない。データ点がある分布に従っていると仮定して、その分布を特徴づけるパラメータを推定することにも使われる。その分布あるいは確率が分からないから、最初、分布やそのパラメータ値を仮定し、与えられた情報からそのパラメータ値を推定し、更新するという使い方ができる。

実際のデータの例を図２に示そう。

図２　実データの分布の例



データはばらつきを持ち、個々の値ではなく確率的な分布として理解される。（文科省作成の中学生成績モデルデータから作成。横軸は数学、縦軸は国語の成績。

丸印（〇）は男子生徒、バツ印（×）は女子生徒。）

統計データを使用する場合、最初に与えられた情報がそのままの形で理解し易いとは限らない。再び子犬のデータについていえば、「大きさ」と「可愛さ」はそれぞれ独立な要因と考えるよりも、例えば（大きさ）×0.3＋（可愛さ）×0.7を要因とした方が、データを整理するときに見易くなる。このような新たな変数を**潜在変数**という。適切な潜在変数を見つけることができれば、変数の数を4から3に減らすことができる。これは統計解析における**「回帰分析」「主成分分析」**等の手法であり、線形代数が効果的に用いられる。図２の成績データでみると、回帰分析の結果、男女に差はみられないことも定量的に分かる。

文章をAIに直してもらうと個性が薄れたと感じることがある。AIが学習データにもとづいて、最大多数が書く標準的表現ーすなわち統計的に最も有りそうな（平均的＝普通の人が書く）表現ーを選び取るからである。

“AI は正しい結果を与える”と見えるのは、上の説明のように、「最も確からしい結果」を選ぶ仕組みが用いられているからである。すなわち「AIの正しさ＝**多数が正しいと思う解の選択**」である。

統計学という数学は、データに基づき、中にある不確実さを「確率」という数字を使って定量化し、その出来事（事象）の**より確かな分布**を推定することを可能にする。

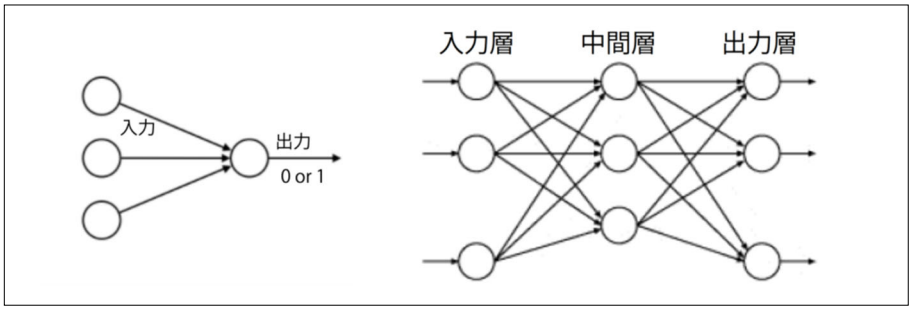
（３）最適化（学習のための数学）

次に、この判断の精度をどのように高めるかを考える。

「AIが学習する」というと、人間のように理解を深めていく過程を想像しがちである。とりわけ「機械学習（Machine Learning）」とか「深層学習（Deep Learning）」という言葉は、そのような印象を強める。しかし、実際には、AIの学習とは、予測と正解（もっともらしい解）のずれ（誤差）を小さくするように内部の調整値（パラメータ）を修正していく過程である。

AIの学習過程を見てみよう。（ここは読み飛ばしてもよい。）

図３　脳のニューロンを模した人工パーセプトロン（左）とニューラルネットワーク（右）



入力の重み付き和に基づいて判定が行われ、これを組み合わせることでニューラルネットワークとして構造を持つモデルが構成される。

ニューラルネットワークなどのモデルは多段構造を持つ。各データに対して予測を行い、その誤差を評価し、誤差が小さくなるように調整を繰り返す。この操作を多数のデータに対して実行することで、モデルは全体として誤差が小さくなる方向に安定していく。ニューラルネットワークの構造はデータをどのように表現するかという側面も持っている。近年広く用いられているTransformerと呼ばれるモデルもその一つであり、自然言語処理の分野で画期的な性能を示した。

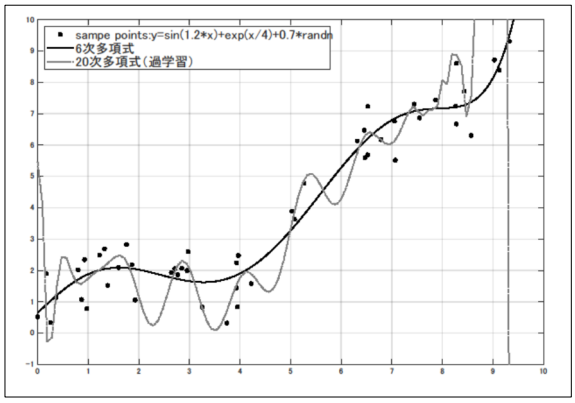
しかしこの過程は常に最適な結果を保証するものではない。「AIの学習」とはAIが自ら理解を深めるというのではなく、誤差（もっともらしい解との違い）を減らすよう設計された調整過程が実行されている過程であると理解す

るのが適切である。

図４　過学習：標本点は

$$y = \sin\frac{x}{2} + e^{x/4} + 0.7 \times (\text{標準正規分布する乱数})$$

で作成



実線は標本点の分布を６**次多項式**で最適化したもの。鎖線は**２０次多項式**による最適化で、乱数で散らばした“揺らぎ”部分を再現しようとした「過学習」の振る舞いを示している。

「学習」の過程では、誤差の調整を繰り返すうちに、結果が不安定になることがある。誤差の蓄積、学習データの偏りや量の不足、あるいはモデルの複雑さ（ニューラルネットワークの層が多すぎる）などが原因である。これを**過学習（Overfitting）**という。AIには、過学習に陥らないように解の最適化を決める過程が組み込まれている。AIの大きな発展は、様々な手法（自動微分や逆伝搬法など）により「学習」の繰り返しを安定に実行できる技術によって支えられている。それが、さらに多層な構造を持ったモデルを可能にした。一方ではモデルの複雑さは、内部の動作の解析を難しくし、AIがブラックボックス化する要因にもなっている。

“AIは自ら学習している”ように見えるのは、複雑なモデルが可能になったことの成果であるとともに、AIがブラックボックスとなってしまい、実際に何をしているのか分からなくなったためである。

（４）エントロピーと尤度（限界と効率のための数学）

AIは万能ではなく、データに含まれる情報の量や不確実性、さらに計算資源の制約によってその性能が制限される。（３）までは誤差（不確実さ）を、「得られた解」と「より確からしい解」との差としてとらえ、その量を小さくすることを考え、その数学を説明してきた。これをより一般的にみると、**エントロピー（不確実さの指標）**あるいは損失関数の最小化、そして**尤度（ゆうど）**を指標として解を選択するという過程であった。尤度とは「どこで止めれば最ももっともらしい結果（最大尤度）が得られるか」を与える判定基準である。

“AIが考えている”ように見えるとき、AIの内部では**損失関数を最小化**するようにパラメータを更新し、その結果として解（出力）を生成している。

（５）数学の実装が受ける制約（AIを支える工学技術）

数学的理論は、計算機上で実装されて初めて機能する。しかしそこには計算資源や通信、エネルギーの工学的制約が存在する。

大規模モデルでは、データや計算量の増大のため、プロセッサ（処理装置）の微細化が試みられる。またデータを上手に個々のプロセッサに割り付けてやらないと、1つのプロセッサに割り当てられた計算が終わっても、他のプロセッサでの計算が終わらず、処理が終わったプロセッサは手持ち無沙汰に待たねばならない。このような結果、計算そのものよりも、プロセッサ間の**負荷の均等化**やデータの受け渡し（通信）がAIの性能を左右する。そのため並列計算や専用ハードウェア（GPU[注４]など）が用いられる。

さらに、プロセッサを繋ぐ（あるいはプロセッサー内の）導線も発熱するので、省電力化（より小さい、あるいは導線部分も短くするような構造、あるいは冷却システムの効率化や、「軽量」なAIモデルの開発）などの工学的課題解決への努力は、AIの発展に必須である。特に省電力はAI産業にとっても最重要課題である。三菱総研の2024年発表の資料によると[文献３]、「生成AIの「社会浸透」と「基盤モデルの大規模化」という2つの要因により、2040年の日本の総計算量は2020年比で最大十万倍以上に達する可能性がある」という。「省電力化した分だけしかAIのための電力使用増をしてはいけない」と

いうような国際的な取り決めがいずれ必要になるかもしれない。

３．AI時代における数理教育

前節で述べたように、AIの仕組みを理解するためには、(１)数値による表現、(２)確率的な不確実性の扱い、そして(３)誤差を減らす調整過程についての理解が重要となる。これらはAIを適切に活用するうえでも重要な視点である。しかし、従来の数学教育において必ずしもこれらが明示的に意識されてきたとは言い難い。今後は、AI時代に向けて、数学に基づいた考え方(ここでは広い意味で**数理**と呼ぶ)を基礎的素養としてどのように位置づけ、教育の中で培っていくかを検討する必要がある。

従来の数学教育では、
（１）定理の証明を中心とした体系的理解、
（２）手計算による計算技能の習得、
が重視されてきた。しかし、AIの基盤となる数理の観点からは、これらの位置づけには再検討の余地がある。すなわち、これらを学習の出発点とするのではなく、その役割や順序を見直す必要がある。

具体的には、
（１）抽象的な証明に先立ち、具体的なデータや例を通じて構造を理解すること、
（２）手計算による処理よりも、大規模データに対する数値計算を通じて現象を把握すること、
を教育の中心に据える必要がある。

証明や手計算は依然として重要であるが、それらは**数理的構造の理解**を支える手段として位置づけ直されるべきである。本節では、これらの観点を踏まえ、①第２節で述べた数理との対応、②教育で留意すべき点、③到達目標の設定、という３つの視点から具体的に検討する。

さらにまた、カリキュラム編成の上で、従来の教育では学問分野を網羅的に扱うことが重視されがちであった。それに対して本稿で強調したいのは、**到達点を明確**にし、**要所に目標を設定**した教育の重要性である。細部の完全な理解にこだわらずに、まず全体の構造と到達点を示し、学習の流れを把握させることが、理解を深めるうえで有効であると考える。本稿では、AIの基盤となる数理的な考え方を教育の対象として明確にし、その「**教育における配置と構造**」を整理して図５～７に示す。

（１）表現のための数理（線形代数）

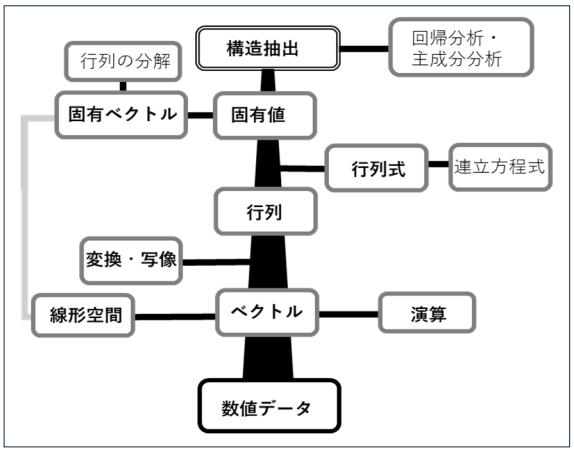
①様々な情報（画像、文章、統計データ）を大規模な数値データとして表現し解析するとき、「線形代数」が重要な役割を果たしている。

情報を多次元空間のベクトルとして扱うことにより、それらをより客観的に扱うことが可能になるとともに、相互の関係を明確に定義できる。「学生、生徒の健康、学力などの統計データ」「日本各地の気象統計」などは関心も高くかつ容易に手に入るデータであるから、大規模数値データの取り扱いを一般論としてではなく、具体的な日常の関心事として取り扱うことができる。具体的問題を扱うとき大規模行列は不可欠であるが、これまでの数学教育では、手計算可能な小さな行列あるいはガウスの掃き出し法の理解までに限られ、それ以降は抽象的なベクトル空間として議論された。

②具体例として、一つのデータと別のデータの関係を「**線形写像**」という性質で示すことができれば、具体的なデータの比較という立場に新しい視点を加えることができる。

また、具体例をもとに、多数の原因因子の相互の関係を見ることができる。与えられた

図５　線形代数における数理構造
：学習ターゲットはデータからの「構造抽出」である



原因因子の軸のままでデータの分布を見るよりも、少し傾けた方が（回転した方が）データ分布の性質が際立って見えてくる。このことを、実データで示してみたい。これで線形写像、1次変換が視覚的に理解できよう。

③これだけの経験をもとに、一般的な方法としての**回帰分析・主成分分析**に導くことは容易になるだろう。限られた講義時間の中では、

[数値データ]→[ベクトル]→[変換]
→[回帰分析・主成分分析]

という一本の流れに沿って学ぶことが全体のストーリーを作るうえで重要である。その中で、数学として線形代数を学ぶために、諸項目を配置すれば、より広い内容を効率的に扱うことができる。

こうして知識の要点となる一本の筋を明確にして、その一本の幹に重要な事項をいくつかの枝葉のように配置すれば、新しい体系的カリキュラムを構築することができるのではないだろうか。到達点を設定した上での、事項の再整理という視点で、学習過程を再構築したい。

（2）確率事象のための数理（統計学）

世の中の（自然現象も社会現象も）多くの現象は、「偶然」という要素を避けて通ることはできない。

①「**偶然**」とは何か、偶然を客観的にとらえるために人間が考え出したのが「**確率**」という数値的表現である。

現象は（「我々の宇宙の創成」や「現在の地球上の生命に繋がる生命の誕生」を除いて）1回しか起きないということはなく、何度も繰り返される。様々な現象（実験）は、丁寧にやっても全く同じ結果を得るのは難しい。現象は確率的なばらつきを持つものとして理解することが重要となる。

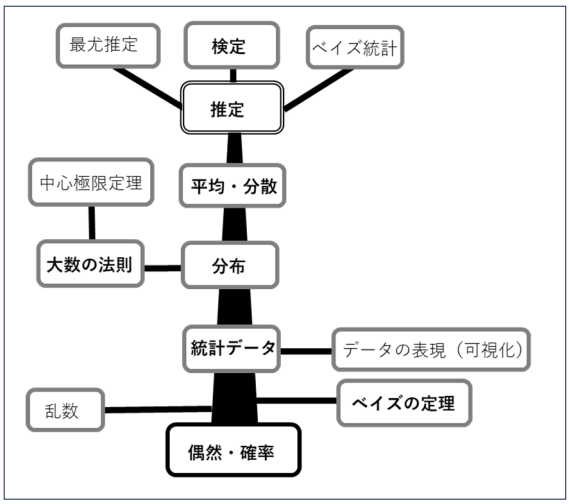
コイントスで表が出るか裏が出るかはもとより、人口における男女比など多くの社会現象は確率的な性質を持っている。さらには、確率ありきで議論を構築する現代の方法論を学ぶ糸口になる。予測可能かどうかという視点が確率・統計の入り口となるだろう。何回サイコロを振れば、各目の出方が1/6に近づくのか、まずは実際にサイコロを振って実験させてみたい。

②偶然に従う数の列を作るにはどうすればよいか（**乱数と疑似乱数**）、また乱数の系列は唯一ではないということから、多くの問題で解が一つではないということ、同じことを繰り返しても、結果が違うということが分かってくるだろう。「メンデルの遺伝の法則の研究」の実際を振り返るのは、良い例題となるだろう。従来の数学教育では一つの正解を求める形式が中心であるが、現実の問題では複数の可能な結果があり、その中でどれがより妥当であるかを評価する必要がある。多くの人々が、「数学は解が一つ」と理解していると思うが、「不確定」なものを扱う数学の体系は新鮮な驚きであるかもしれない。

確率はもともと決まっているものなのか（**事前確率と事後確率**）ということから、**ベイズの定理**を学ぶことは必須であるが、「**ベイズ統計**」に本格的に踏み込むのは難しい。「**最尤推定**」をきちんと講義するのは難しいが、**エントロピー**や**情報量**という言葉はしばしば耳にするようになってきたので、その意味はしっかり学んでおくのが良い。

③重要となるのが、「もっともらしさ」をどのように評価するかという視点である。仮説のもとでデータがどの程度整合的であるかを考えるという枠組みは、統計的な「推定」や「検定」の基礎となっている。

図６　統計における数理構造
：学習ターゲットである「推定」を中心として、最尤推定、仮説検定、ベイズ統計などの方法が展開される



限られた講義時間の中では
[偶然・確率]→[平均]と[分散]→
[分布、正規分布]→もっともらしさ（**尤度**）というような一本道にそって確率･統計を学ぶ必要がある。これらに対して、**誤差とは何か**を考えること、「最小二乗法」や「中心極限定理」あるいは「推定・検定」などの話題が豊かな枝葉となってくれるだろう。**ベイズ更新**がAIの基本にあり、最適化という問題の入り口になるからである。統計学の中で、微積分と線形代数が主要な技法の一つであることを度々経験するだろう。

（3）最適化という数理（最適化法）

我々が日常の問題の解を見つけようとするとき、「**試行錯誤**」という方法をとることは多い。この考え方は自然であり、重要なものである。

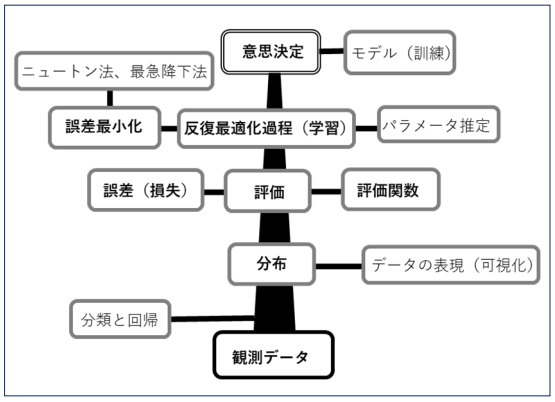
山の上から麓を目指すとき、①最も急な斜面に沿って下る（**最急降下法、勾配降下法**）という方法がとられることは多い。一般の場合、山の高さに対応する指標（等高線、評価関数、目的関数）を考えることができれば、問題はより明確になる。「徐々に解に近づく」という考え方は、従来の数学教育の中での一度で正解を求める問題形式とは異なるものであるから、明示的に扱う必要がある。

②正解に近づくに従い、局所的に平坦な場所にとどまってしまう可能性（**局所最適解**）があり、また**大域的最適解**の近くの広い範囲にわたって指標の変化が小さくなるなど、様々な現象が現れることがある。これらの結果、最急降下法だけでは解の探索が遅く困難になる。それに対して種々の加速法が試みられている。大域的最適解に到達できるという保証がある場合には、落ち着いて正解を探すことができるし、その価値も明確になる。一方で、数値的な取扱いは、誤差が蓄積されるから試行回数が多ければ良いというわけでもない。**最尤推定**は、統計モデルのパラメータを決定する基準の一つである。最適化問題の具体例としては、物理で用いられる**モンテカルロ・シミュレーション**、あるいは**画像修復**などは分かり易いだろう。AIの具体的な仕組みに踏み込むことは数学学習としては適切とは言えないが、近似計算の回数を重ねても不安定にならないということが根本にあるのだということは、話題として取り上げることはできるかもしれない。

③具体的な最適解探査の方法は、個々の問題によって多様である。現在の状態から少しずつ改善していく勾配降下法のような方法は、この考え方を具体的に示すものである。

例えば
ニュートン法→最適解の存在と実行可能性
（凸最適化問題、線形計画法、非線形最適化）
といった流れの中で、代表的な加速法の考え方を示すことができる。最適化法というのは、従来の教育の中では、多くは取り上げられなかった問題であろう。時間が許せば、**時系列解析**の問題にも、足を踏み出すきっかけ程度には触れてみたい。

図７　機械学習における数理構造
：学習ターゲットを「意思決定」とし、誤差最小化に基づく反復最適化過程によってこれを実現する



（4）数理的考え方を支える実データと計算環境（大規模データから学ぶために）

（1）～（3）の考え方を実際に身につけ役立たせるためには、公的機関（政府統計や世界銀行など）が提供する実データを用いた学習が重要ではないだろうか。実データは、それだけで学習者の新しい興味や意欲をかきたてる可能性が高い。それ以上に、「信頼できるデータ」の価値ということも学ぶべき問題である。
そのようなデータを教材として用いると、①従来行われてきたような、“紙と鉛筆”で学ぶ線形代数や“作られた少数のモデルデータ”での

統計学では済まされなくなるだろう。しかし初めて生きた数学を経験することができる。

②さらには“**分布**とは”、“**ノイズ**とは”、“**誤差**とは”というように具体的な視覚的概念として目の前に現れてくる。

③実問題に対しての計算を行うということは、**数理の様々な分野の横断的な連結**を促す可能性が高い。これまで繰り返し述べてきたが、主成分分析は線形代数のなかの座標変換という問題とともに、主たるベクトル空間、**固有値の分布と特異値分解**というこれまで教育の中であまり触れられてこなかった問題への糸口を与えることができる。

AI時代の（大学）数理教育は、理工系学部のみならず人文社会系学部において「**演習**」をどうするかという大きな問題にかかわってくる。理工系学部では数学の講義には必ず演習が付随している。学生数が増えた結果、演習の担当は教員スタッフだけでは足らず、大学院生をTA（Teaching Assistant）として雇用して担当してもらっている。これは単に人手不足のための対応ではなく、大学院生の教育の一環として、また経済基盤の整備・安定化として重要であると考えられている。TAの経験は教育経験として評価される。

一方、文系では、大学院学生の絶対数が少ないこととともに既存の教育のなかで数理科学的基本が培われていないことにより、絶対的な人手不足の状態にある。これまでは、人文社会系の数学教育は、あまり必要を感じていなかった人が多いこともあり、教養として学ぶかあるいは理工系の数学を単純にスケールダウン（簡単化）して行うという色合いが強かった。人文社会系人材が主体的にAIを使えるようにするということを、教育の目標に明示的に掲げる必要があるだろう（**教養から実務へ**）。

数理・情報教育のカリキュラムを体系化し全学的に展開するために、2017年に東京大学の中に**数理情報教育研究センター**が設立された。筆者たちは、経済学部の中に文科系専門学部生のための数学講義を開設して担当してきた。そこでは従来型の演習の代わりに、PCによる計算や描画を通して、「**『計算＝実験的方法』により数学を学ぶ**」という試みを行ってきた。計算アプリとしてはMATLABを用いることにした。[注5)] 講義で使用するプログラムは全て予め教員側で用意し、学習者が自由に使い書き換えができるような形で提供した。[注6)]

数理教育は微分積分を前提にしなくてはならない。微分積分学と線形代数とは大学教養課程で選択科目として講義されるのが標準であったが、これらの内容や重要性も、これまでとは違っていくだろう。

４．おわりに

線形代数や微積分、統計学、最適化手法といった各分野が、それぞれの体系を持った木として立ちながらも、それらを支える工学技術という共通の地平において、茂った枝を互いに伸ばし、複雑に絡み合っている。この交差点にこそAIの計算技術の本質がある。これらを具体的な例に引きながら、知識が有機的に結合した体系的な教育課程を構築することで、学生はバラバラな数式の中に一本の確かな道筋を見出すことができるはずだ。これらの内容を限られた教育時間の中で実現するためには、教育内容の選択と重点化が不可欠となる。限られた時間の中では、従来重視されてきた内容の一部を精選し、上記の内容に重点を移していくことが求められる。

AI時代を迎えて、数学（数理科学）は専門家だけのものではなくなりつつある。それは、AIを使いこなすための基礎的な「読み書き能力」として、多くの人に求められるものになったからである。

【脚注】

注１）東京大学大学院工学系研究科教授（物理工学専攻）を定年退職の後、大学総合教育研究センターおよび数理・情報教育研究センター特任教授を務め、2022年に退職。

注２）2026年4月時点では、英文 Wikipedia では直っているが、日本語版ではまだ訂正されていない。

注３）自動運転におけるデータ処理はこのようにして行われる。

注４）GPU（Graphics Processing Unit）は高精細な PC ゲームや動画編集のために作られた専用プロセッサ。

注５）MATLAB を選んだ理由は、2019年から東京大学全学で（学部、大学院）学生、教員、研究員、職員の区別なく誰もが使える環境ができたということに加えて、数学との親和性が非常に良く、初学者にとっての障壁が低いからである。実際、講義のはじめに1.5回分くらいの時間をかけてインストールから文法の説明をして、講義と演習に使用することが可能だった。MATLAB はバージョン管理が必要ではなく、AI に関する様々なプログラムを提供している。欠点を挙げれば、有料であるため（バージョン管理が必要ないということはこういうことだ）、社会に出たとき使える環境があるかどうかは分からない。

注６）この講義では、最適問題と統計解析を学ぶことを目標に、そのための線形代数と微分積分に重点を置いた。統計学に十分踏み込むことはしなかった。どのようなテーマを選び、どのようなプログラムを準備したかに興味のある方は、以下を参考にしてほしい。

「データ科学のための微分積分・線形代数」藤原毅夫、藤堂眞治（東京大学出版会、2021）

【参考文献】

文献１）「MATLAB を用いた統計手法」藤原毅夫、島田尚　（東京大学出版会、2025）

文献２）トーマス・ベイズ（1701または1702―1761）はイギリスの長老派牧師で数学者。今日われわれが見るようなベイズの定理にまとめ、さらに発展させたのはピエール＝シモン・ラプラス（1749–1827）。

文献３）三菱総合研究所　2024年8月28日の Web ページに述べられた提言「生成 AI の普及が与える日本の電力需要への影響」