

データサイエンス学部に於ける数学教育とデジタル人材育成

渡瀬泰成[†]

[†] 立正大学データサイエンス学部

[†]ywatase@ss.iiij4u.or.jp

キーワード データサイエンス, 数学教育, ロジック, 証明, 問題解決

1 はじめに

立正大学データサイエンス学部はデジタル時代の人材育成のため 2021 年度より開設学生を迎え本格的に学部が始動した。理系に限らず将来有用となるデジタル人材育成のため文系へも門戸をひろく学生を応募し入試で数学を受験科目にかならずしも必須としていない選抜区分があるのも特徴といえる。筆者は 22 年度より数学演習科目（微積分学演習，線形代数学演習，統計学 II 実習，いずれも後期 2 コマ続きの演習）を担当した。データサイエンス学部に於ける数学教育内容について僅かな経験であるが将来データサイエンティストとして期待される人材が身に付けるべき数学の基礎力について若干の考察を述べる。

2 演習実習科目について

24 年度後期に担当した「統計学実習」と「線形代数演習」を例にとり高等学校で数学 III まで履修し工学・理学を目指す所謂理系志望ではない学生への数学教育の一端，取り組みについて紹介する。高等数学の中核をなす微積分学演習，線形代数学演習については文献 [8] に於いて詳しく解説されているので参考されたい。

2.1 統計学実習 II

統計学実習 II は後期に 2 クラスに対して実習を担当した受講者数は各 36 名であった。実習の方針は統計学の概念と基礎的内容の理解，統計ソフトのプログラミング入門であり最低限の目標は基本的な定義の理解，確率分布の理解と統計量の推定の方法を理解説明発表できること，及び統計演習の内容に対応した R 言語でのプログラムの実装であった。実習は基本的な演習問題に取り組み 1 題 20 分程度考えさせノートに書いて解かせる。よって 90 分で 3 題程度解くことになり概念の習熟を図っている。問題を解いている間は教室を歩いて一人ずつ答案

を見て回り途中でつまづいている場合はアドバイスを与える指導を行った。個々人に対応すると真剣に課題に取り組む割合が増すように見て取られた。これは数学を勉強したい意思の表れと分析している。PC 利用のハンズオントレーニングとして文献 [1] を参考に統計ソフトウェア「R」（以下単に R と表記する。）の描画とコマンドの初級レベルを実習した。具体的な実習内容は以下の表の通りである。

表 1. 統計学実習 II の内容

講義	内容
第 1 回	R/RStudio の導入, R でのデータの扱い, サンプルのロード, コマンドの例.
第 2 回	R による宅地価格・台風上陸回数データの分析, ヒストグラムと統計量.
第 3 回	散布図と相関関係 民間・公務員の給与データを R により相関を分析.
第 4 回	相関関係 テスト結果からの科目間の相関.
第 5 回	条件確率の復習 (ベイズの定理) メレの逆説の R によるシミュレーション描画.
第 6 回	離散確率分布 (1) ベルヌーイ試行 二項分布の性質.
第 7 回	離散確率分布 (2) 2 項分布 vs ポワソン分布, 相対頻度の極限のシミュレーション.
第 8 回	連続確率分布 正規分布の密度関数, R によるコーディング.
第 9 回	連続確率分布 長期/時系列データ (TOPIX) の分析, 対数尺度, 平均収益率と R によるコーディング.
第 10 回	推測統計: 分布の精密化, 確率変数の導入.
第 11 回	期待値と分散の線形性の計算と証明.
第 12 回	演習: 母平均の推定.
第 13 回	演習: 不偏分散の導入, t-分布.
第 14 回	演習: 母分散の推定, χ^2 -分布, R による描画.
第 15 回	課題の検討.

2.1.1 半期の演習を終えて

本実習の準備段階では演習問題を黒板で解いて他の学生に内容を説明するという想定していたが受講生を観察して黒板は各自のノート，他者へに説明は個々人

への質疑応答という形に変容してしまった。Rは英語（シングルバイト）のソフトで日本語環境でも動作するがダブルバイト文字がコマンドに混入するとバクとなり扱いづらい面がある。この手の泥臭いデバッグを重ねないとスキルは向上しないのだがあまりに時間をとられるので事前に配布資料を用意してそこからスクリプトをコピー&ペーストすることも実施しRの機能を実習した。R言語が得意であることが実務で求められることは無いかもしれないが、未知のコンピュータ言語を独力で使えるようになる訓練にはなったと思う。

2.1.2 受講学生の分析

演習授業では1人ずつ演習の進捗を見て分からない箇所を指導した。計算問題は例題を追いながらフォローできる。証明問題には苦手意識あり。Rのコーディングの実習は熱心に取り組むと言ったことが15回の授業を終えての感想である。ベルヌーイ試行を定義してベルヌーイ分布を教えた場合と数学用語に戸惑う様子であり数式だけの説明では足りないので硬貨投げやサイコロでの例で離散分布をつくり期待値の計算を行うと理解が進むのである。新たな概念の理解は簡単な例や道具を利用したり実験しないと獲得できないものであろう。確率の計算で二項定理 $((x+y)^n$ の展開公式)の証明の方針を質問しても帰納法を使うと答えは得られない。証明してみようと色々実験すればいいのと思うので $n=1,2,\dots$ と実験をデモすることもあった。理系文系の区別は意識せず授業をしているが文系の学生が多いのかと推察している。ここで定量的なデータを見てみる。証明についてどの程度理解しているか任意に提出するレポートとして小テストを実施した。内容は三段論法を解説を与えた後に幾何の知識を調べる課題(問題1と問題2)と問題3において三段論法を応用を目論む課題である。図形が合同ならば大きさ(面積)も等しいという公理は明示していない。また面積の定義も明示していない設定で与えた多少意地悪な問題である。

問題 1 三角形の合同条件を述べよ。この問題は調べて解答してよい。ただし調べた出典も記すこと。(この三角形の合同条件はユークリッド幾何の公理である。公理とはいつでも前提としてよい論理式である。)

問題 2 ヘロンの面積の公式を述べよ.. この問題は調べて解答してよい。ただし調べた出典も記すこと。

問題 3 二つの三角形ABCとDEFの辺の長さで $AB=DE$, $BC=EF$, $CA=FD$ ならば三角形ABCの面積

と三角形DEFの面積は等しいことを証明せよ。

36名中30名の解答を得た。問題3の解答は以下の3つに分かパターンに分類される。

パターン 1 二つの三角形が合同であることより面積も等しいと結論。(解答者全体の40%)

パターン 2 三辺が等しいのでヘロンの面積の公式を適用すると面積は等しいと結論。(36%)

パターン 3 二つの三角形は合同である。対応する辺と高さも等しいので面積も等しいと結論。(7%)
(残りは合同であるとしか解答しなかった。)

この問題の正解は難しい。パターン1の場合は「合同ならば面積は等しい」が公理であると調べる必要がある。または別公理から導出する必要がある。パターン2の場合はヘロンの公式を認めるところから出発するが、関数への等しい数を変数に代入するとき関数は等しいという代入の導出スキーム(型)が必要である。または実数の公理から等式を導くことが必要である。パターン3は面積をテンポラリーに定義して推論しているため論理的推論の努力と創造性の跡が分かる。総じて証明が苦手とか推論力が不足しているとはいえない。

「証明は苦手」という意識を払拭して勉強すればよいであろう。今後の学業の糧となるようこの課題の最後に学生へ「数学が分からないとは」と題して以下を補足した。曰く「数学の特定の理論(Theory)や問題や証明(Proof)が分からないと思うことはしばしば起こる。この原因は理論で定義した用語を正しく理解していない。成り立つとして使っている前提(Premise)、公理(Axiom)に気付かないことであったり基本的な定理(Theorem)が見えていないこと、証明や説明に現れる推論規則(Inference Rule)が見えていないことに起因する。時間を掛けて紐解いて考えて行けば必ず分かることである。数学は様々な理論(集合論、実数論、位相空間論、群論,...等々)から成り立つ。理論を構成する要素は、1. その理論を規定する公理 2. 固有の記号とか単語(点とか面等)そして3. 推論規則でありこれら1-3をまとめてその理論の言語(Language)という。数学を理解するのは外国語学習に似ている。基礎となる文法(ロジック)を身に付けて単語を正しく覚えて言葉を運用することに似ている。継続して取り組んで諦めず学習してもらいたい。」と励ましている。

2.2 線形代数学演習

統計学以外の教科として線形代数学演習について言及しておく。対象学年は2年生である受講生は22名であり

半数は再履修者という状況。本演習では数式処理ソフトウェア Maple を利用しつつ線形代数を学習する。Maple はグレーブナ基底の計算するため利用した時期があったので行列計算、記号の多い数式の簡約に威力を発揮することは心得ていたが線形代数に適用するのはあまり想定していおらず参考書 [3, 4] を参考に演習に取り組んだ。受講生は掃き出し法は手計算で出来るレベルなので行列計算の練習などは特に問題ないレベルであった。平面ベクトルの話では三角関数を使う場面があるとつまづきを見せる。授業中に三角関数は得意か不得意か調査し三角関数の復習中心の演習を交えた。またプロシジャーにより行列を定義しプログラムから呼び出す単純な演習があったが「サブルーチン」は初めて接したとの指摘があった。線形空間の公理、空間の生成系、基底等の定義は抽象的なのか理解しにくい様子もあったので抽象的な空間は扱わず数ベクトル空間での説明した。全員 Maple ははじめて利用するのでマニュアルを解説しながら基本操作から練習し少しずつコマンドと数学的な対象、ベクトルや行列の入力を実習した。以下の表が演習した内容となる。2コマ続きの演習であったので前半のコマで基本的な問題(行列やベクトルの計算問題/線形代数の理論的な問題)を紙とペンを使い実施、後半のコマでは Maple 向けの話題や前半コマの計算問題を Maple で検証するものであった。

表 2. 線形代数演習の内容

講義	内容
第 1 回	集合、写像、数・Maple の使い方・計算.
第 2 回	連立一次方程式の解法、消去法/Maple のマニュアルを読む
第 3 回	行列、一次方程式/Maple での行列の表現と使い方・計算
第 4 回	掃き出し法の手計算・逆行列の演習/Maple での行列の操作.
第 5 回	ガウスの消去法・手計算と Maple. 掃き出し法を Maple で書き下す
第 6 回	行列式とクラメールの公式演習
第 7 回	線形空間の定義や性質/連立方程式と線形写像の対比
第 8 回	基底とランクの計算演習・Maple で確かめる
第 9 回	部分空間と線形写像の解空間・次元定理/Maple.
第 10 回	Maple で関数定義: 特殊な行列 (Hilbert 行列) を生成するなど.
第 11 回	Maple 演習, ベクトルの内積と外積, 直交化とその幾何的意味.
第 12 回	三角関数の復習, 余弦定理/Maple 演習内積を変形して面積求め式を簡約=外積
第 13 回	回転行列と加法公式群を導出.

講義	内容
第 14 回	線形写像の表現. 定義や講義の復習
第 15 回	課題の検討.

2.2.1 半期の演習を終えて

線形代数は一次方程式系の解と線形写像の像と余像の関係を抽象数学のよい題材である。この方面は理解度は期末のレポートの出来から 50 %程度であった。ベクトルと図形の関係は解析幾何の要素もあり目でみて考えられるので受講生も一生懸命取り組んでいた。三角関数を復習したあと 2 つの平面ベクトルの外積がベクトルが囲む平行四辺形の面積に等しいことの証明を演習に出したが 2 名ほど正解, 6 割程度は方針は正しくできていた。高校数学も別の視点から眺め直してみると今度こそは理解するぞとの意気込みがあった。反省面としては行列のユニタリ行列による類別のはなしや線形計画法のはなしが出来なかった事であろう。最適化の手法で別分野であるが線形不等式は応用は広いので触れるべきで Excel のアドインソフト Solver¹で例題を解く程度は含めてもよかった。

2.3 演習科目の意義

第一には講義科目の内容の復習と理解の定着にあらう。統計学と線形代数の実習/演習で問題に取り組む受講生に接して以下の事項が重要と考えた。

- 容易な問題 (用語の説明させる問題, 単純な計算問題) から始める。これは演習に参加するモチベーションを維持する。成功からつぎの挑戦する意欲を引き出す。
- なるべく図を交えて問題を説明する。考えのきっかけ, ヒントを与え 10 分程度自力で考えてもらい個々の学生の進捗をチェックしつつアドバイスを与える。
- 分からないことは遡って中学/高校数学の内容を説明する。
- 証明が要するところは出来そうならやらせる。推論のステップを曖昧にしないで考えるように指導する。(一緒に考える)
- 考えを述べたり, 質問を述べるよう発言を促す。

これらは学生が自ら考え推論しながら問題解決するという経験を積める為の配慮であった。問題解決の能力が向上し「どうすればいいのか」といつも考えて解決策を提案できるようになってほしいからである。演習を通じてプロブレムソルバー (問題解決者) になってほしいのである。他方次の側面も大事であろう。即ち数学の抽象的概念を理解する訓練である。幼い子供はおもちゃの飛行

¹Solver は線形/非線形問題の最適解を求める強力なツールである。

機や車で遊ぶことによりその概念を獲得しているように見える.[9] 数学の概念の獲得も演習の時間に簡単な例(おもちゃに相当する)をいろいろいじってみたり疑問点を質問したりして少しでも獲得出来るるとよい. 概念の取っ掛かりでも掴めると理解も随分進むからである. 従って演習担当は以上の教育的配慮点を踏まえて演習問題や題材を準備する. 或いは演習内で補足できるとよい. 演習を通して理解力, 概念の把握力, 問題解決力のポテンシャルが上がることを期待したいのである.

3 デジタル人材育成に向けて

データサイエンティストに期待されるスキルを支える数学技能に必要な要素について実習での学生を観察して考察した内容を論ずる. 経済産業省のデジタルスキル標準(DSS) [11]によれば「データサイエンティスト」は3つに類型化している.

表 3. データサイエンティストの種類とその役割

役割区分	役割
データビジネスストラテジスト	ビジネスに精通して事業戦略に基づくデータ活用戦略を立案する役割
データサイエンスプロフェッショナル	データサイエンス領域の専門性に基づき、データの処理・解析や、その結果の評価等を担う。
データエンジニア	データ分析環境の基盤の構築運用を担う。

本学の教育では2番目のデータサイエンスプロフェッショナルの要求を満たすべく授業が組まれていると考える. このロールに必要なスキルセットには1. 数理統計・多変量解析データ可視化, 2. 機械学習・深層学習が挙げられている. 同標準でこのこの区分に期待するは役割は「データサイエンスプロフェッショナル」は、データの処理・解析を行うほか、その結果を評価し、新規事業の創出や現場業務の変革・改善につながる知見を生み出す役割を担う。また、現場部門でのデータ活用の仕組みづくりやエンドユーザーに対する教育・サポートを行うという役割も担っており、データの処理・解析だけではなく、その結果の活用の場面においても一定の責任を負っている。」([11] p110 より引用)である. 名称は「データサイエンスプロフェッショナル」と新規に見えるが期待される役割は従来のIT部門のアナリストの役割に近いと言える. 技術は日進月歩に進歩するようで存外コアスキルは変わっていないと考える. スキルセットの項目2. 機械学習・深層学習については1950年代のサイバネティクス全盛期(パーセプトロンの研究, 列車の自動運転への応用)の人工知能の研究から原理的に変わらない. 1980年後半90

年代前半にもAIは盛んに研究された. 今般の計算機能力の向上に伴うAI技術革新の期待による第三期といってよい. 根本的な概念は70年前に成立した科学領域の延長である. 数学の基礎力(数学の理解力と言ってよであろう)をしっかり身につけていけば最新の技術に対応できるといえよう.

3.1 ロジックの重要性

演習の授業から垣間見える証明への取り組みや数学理論を理解するためのロジックについて論じたがプログラミングの教育でも証明の技術は重要である. プログラムを作成する場合プログラムの論理的な正当性と停止性を検証しなければいけない. これはプログラミングの基本である[2]. プログラムの教育で基本的なアルゴリズム, 互除法や行列の計算のアルゴリズムの計算量とプログラム検証は基礎的な内容でありこれもまた数学の証明と同様に厳密に推論しないと導けない. 定理証明によるプログラム検証やモデル検証4-5年前は先端技術と考えられていたし[5, 6], 機械学習により自動証明も研究されている[10]. 半導体の回路設計の検証は20年以上定理証明の手法が用いられロジックによる検証/証明は重要な技術である. 卓近な例でいえばアプリケーション開発に於いての仕様書作りでも論理的に整理されていないとプログラム開発が困難となる. プレゼンテーションも然りであって論理的に整理されていない発表は理解されない.

3.2 サイエンスマインド

科学的であるとは何であるのか改めて思うと筆者が学部時代学んだ「科学基礎論」の講義を思い出す.[7]「科学で重要なことはその方法論を学ぶことである」と教わった. 実験を繰り返しデータを集め, 仮説を立て数学を用いて厳密に証明する方法である. 概念を理解したり獲得するには色々実験を重ねること, これは思考実験であっても紙面上に色々作図することでもよい考えながら自分の手を動かして理解するまで地道な努力を惜しまないことが要求される. 数学の勉強をするのであれば記号論理の推論規則, 算術の帰納法, 述語論理の量子子の初歩を使えるようにしてから演習問題に取り組んだり具体例を考えて思考実験をする. 分からなければ証明してみようと前向きな実験するマインドが肝要. 考えたことを言葉にして同級生と話したり文書にすることも考えの整理には大いに役立つ. 前述したように実習の授業では自分の考えをまとめて発表する訓練ができなかったのは反省点であり改善点である. 学生を消極的から積極的に発言す

るよう導く工夫は課題として残る。実習のうち何回かは周りに座っている学生同士相談も可としグループ作業で課題を考えて解答してグループで発表する方法も考慮したい。

4 まとめ

ロジックを使って推論する術を身に付け問題に挑戦する気力が備われば独力で考えて問題解決する準備は完了する。学校では演習やゼミで問題解決に取り組み訓練を積む。社会に出れば職場の同僚と問題解決に取り組みデータサイエンティストとして成長し活躍する。このようなキャリアパスもあっていいと考える。

ロジックの分野では前述の「記号論理の推論規則，算術の帰納法，述語論理の量子子の初歩」の知識の習得。数学の基礎知識として必要なものは，集合論から集合の演算，集合上の関係と同値類の考え，位相空間論からは距離の概念を習得すれば努力次第で確率空間までは理解できると思う。リテラシーレベルでは応用数学では LP や最適化理論など知っていればよい。

まずは推論規則による「証明の練習」そして「考えること」，産業界という海原に出る前の曇りの水練である。

参考文献

- [1] 小暮 厚之, R による 統計データ分析入門 (シリーズ統計科学のプラクティス 1), 朝倉書店, 2009.9
- [2] 林 普, プログラム検証論 情報数学講座 (8), 共立出版, 1995
- [3] ユージン・W. ジョンソン (東野 勝治他訳), MAPLE5 線形代数, インターナショナル・トムソン・パブリッシング, 1997.7
- [4] 守谷 両時, Maple で数学を一線形代数編, 海文堂出版, 1996.4
- [5] 萩原学, アフェルト レナルド, CoqSSReflectMathComp による定理証明, 森北出版株式会社, 2019
- [6] 井田哲雄, 浜名 誠, 計算モデル論入門 計算モデル論入門 (Information Science & Engineering F), サイエンス社, 2006.12.1
- [7] 柳瀬 睦男, 物理学の思想と方法, 三省堂, 1978
- [8] 友永 昌治 (立正大学). 文理系データサイエンス学部における数学教育, 国際 ICT 利用研究学会 特別研究会 (第 2 回), 2023.6.25
- [9] 小野 勝次, 思考の枠組 I, 15 巻, 3 号, p. 125–135, 科学基礎論研究, 1980–1982
- [10] Kaliszky, C., Urban, J. Learning-Assisted Automated Reasoning with Flyspeck. J Autom Reasoning 53, 173–213 (2014).
- [11] 経済産業省, デジタルスキル標準 ver.1.1, 2023.8 https://www.meti.go.jp/policy/it_policy/jinzai/skill_standard/