Вісник НАН України. — 2002. — N 5

І. СЕРГІЄНКО, А. ГУПАЛ

ІНДУКТИВНА МАТЕМАТИКА

Для багатьох читачів викладені у цій статті міркування стануть вельми несподіваними: мовляв, чи доречно говорити про недосконалість математики тоді, коли вона досягла, здавалося б, небувалого розквіту? Однак усе не так однозначно. Фахівцям добре відомі всі виявлені в минулому столітті протиріччя, розбіжності й розчарування з приводу основ «найнепорушнішої» з наук. Автори цієї статті привертають увагу до основних моментів, які, на їхній погляд, не розглядалися раніше, і роблять спробу визначити нові шляхи розвитку сучасної математики.

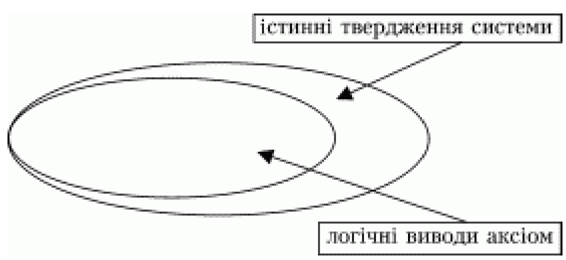
ВІД ДЕДУКТИВНИХ ПІДХОДІВ — ДО ІНДУКТИВНИХ

Для отримання своїх результатів математика протягом багатьох століть використовувала особливий метод — метод дедуктивних висновків з невеликої кількості самоочевидних принципів, званих аксіомами. Цей метод був відомий ще стародавнім грекам, а тепер з ним знайомий кожен школяр.

Аксіоматичний підхід передбачає дедуктивний спосіб міркування: з істинних посилок (аксіом) за допомогою певних правил отримують істинний вивід. Виникає ланцюжок послідовних істинних виводів, одні з яких є аксіомами, а інші логічно випливають з попередніх. Евклід уперше використав аксіоматичний підхід у своїх знаменитих «Началах» геометрії. Арістотель на основі дедуктивного методу побудував формальну логіку. Дедуктивний метод дав змогу математикам «вилучити» з аксіом безліч теорем про числа і фігури.

Аксіоматично-дедуктивний підхід почав інтенсивно розроблятися наприкінці XIX ст., і склався певний стереотип: багатьом здавалося, що в рамках аксіоматичної системи будь-яке істинне для неї твердження явно доказове.

У 1931 р. абсолютно несподівано з'явилася знаменита теорема Геделя про неповноту, яка завдала нищівного удару по аксіоматичному підходу і надіях математиків побудувати дедуктивну математику. У будь-якій аксіоматичній системі, що включає відому нам арифметику цілих чисел, існують істинні твердження, не доказові в рамках даної системи. Тобто логічні наслідки аксіом — це лише підмножина всіх істинних тверджень (див. схему).



Крім того, виникає парадоксальна ситуація: процес додавання скінченної кількості нових аксіом (іншими словами, нових знань) також не робить систему повною. Пізніше фахівці визнали результат Геделя найсильнішим за всю історію математики, хоч і він має явний негативний відтінок. Філософи присвятили теоремі Геделя численні спеціальні монографії, оскільки вона зачіпає питання пізнаванності світу.

По суті, ідея дедуктивного підходу виявилася утопічною: не можна вивести всі істинні твердження на основі лише істинних посилок. Тому відразу ж багато математиків почали висловлювати сумніви щодо цінності такого підходу. У загальних рисах вони зводилися ось до чого.

Завдяки разючому успіху геометрії Евкліда як дедуктивної системи вченим почало здаватися, що наука має будуватися за дедукцією від загальних правил до конкретних прикладів. Це неправильно, особливо якщо йдеться про процес навчання. Наш мозок, мабуть, добре справляється з відтворенням цілісного образу об'єкта за його окремими елементами і з узагальненням найрізноманітніших конкретних уявлень. Побудова ж ланцюжка логічних міркувань — неприродний вид діяльності для людини.

Відомо, що людина не наділена здібностями дуже швидко рахувати, а також проводити логічні побудови. Зате вона чудово розв'язує задачі розпізнавання у просторі і в часі, а найголовніше — має у своєму розпорядженні довершену програму самонавчання. Причому в повсякденному житті ми однаковою мірою спираємося як на «позитивні», так і на «негативні» приклади. Скажімо, у базу даних будь-якої медичної експертної системи обов'язково закладені відомості відносно хворих і здорових пацієнтів. І це зрозуміло, оскільки насправді інформація міститься не тільки в істинних, як це вважалося раніше, а й у хибних твердженнях. Тому не дивно, що всі спроби математиків побудувати повну дедуктивну систему на основі лише істинних тверджень виявилися невдалими.

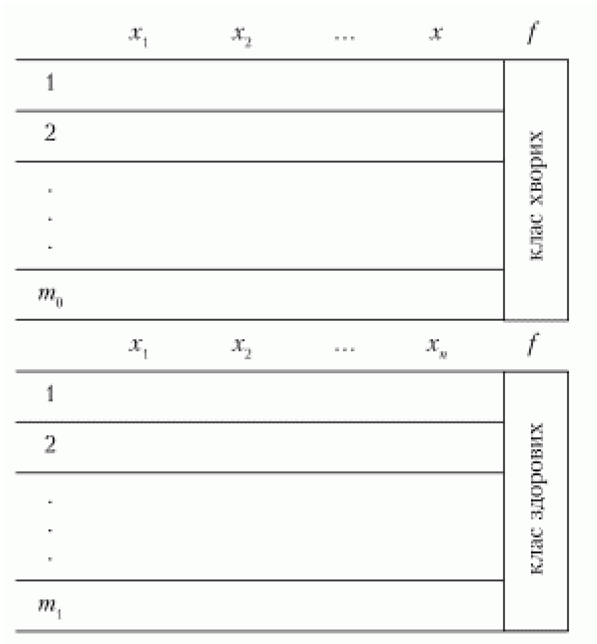
Альтернативою дедуктивному підходу є індуктивний підхід — як спосіб міркувань від окремого до загального. У своїх дослідженнях вчені часто користувалися методом узагальнень. Ньютон виводив свої знамениті закони з експериментів і спостережень. Спробуємо стисло описати цей спосіб.

На практиці ми стикаємося з новими фактами, які не можна пояснити на основі відомих закономірностей, оскільки причини або закони, яким вони підпорядковуються, невідомі. Постає запитання: як на основі емпіричних даних або статистичної інформації щодо описів скінченної кількості об'єктів побудувати процедуру, котра дасть змогу оцінити потрібні нам характеристики будь-якого іншого об'єкта і показати, що достовірність виводу або ефективність цієї процедури зростає із зростанням накопиченої інформації?

У цьому запитанні — суть проблеми індукції. В індуктивних виводах, на відміну від дедуктивних, істинність посилок і правильний логічний хід міркувань не можуть забезпечити повну достовірність виводу. Багато фахівців зазначають, що вся складність проблеми індукції полягає у неможливості сформулювати однозначний алгоритм її розв'язання. Лауреат Нобелівської премії, логік і філософ Б. Рассел заявив: «З часів Лапласа намагалися вивести процедури індуктивного виводу і обґрунтувати їх ефективність. Тепер всіма визнається, що всі ці спроби були безуспішними» [1].

Результати, отримані в останні роки співробітниками Інституту кібернетики ім. В.М.Глушкова НАН України, дають підставу вважати, що проблема індуктивного виводу не така безнадійна, як це здається на перший погляд [2—7]. Виявляється, що сучасну математику цілком можливо будувати за допомогою індуктивного підходу.

Ми вивчаємо природні об'єкти, вимірюючи їхні характеристики за допомогою різних пристроїв або приладів. Для кожного об'єкта, що спостерігається, є його опис у вигляді вектора (x_1, x_2, …, f). Тут x_1, x_2, …, x_n — вимірювання об'єкта, а f — значення характеристики, яка нас цікавить. Візьмемо для прикладу медицину, коли x_i, i = 1, …, n— аналізи, f — стан пацієнта (хворий або здоровий). Часто величини x_i, f називають ознаками. Нехай є m таких описів. Усі вони утворюють так звану навчальну вибірку. Тоді медична експертна система з деякими спрощеннями матиме такий вигляд:



Тут m_0 — розмір класу хворих; m_1 — розмір класу здорових; $m = m_0 + m_1$— розмір навчальної вибірки.

Потрібно побудувати таку процедуру індуктивного виводу, яка за вимірюваними x_1, x_2, …, x_n і навчальною вибіркою дасть змогу визначити характеристику f.

Наступний основний момент: слід навчитися відрізняти об'єкти від їх описів. У повсякденному житті ми часто ототожнюємо об'єкт з його описом, насправді ж це різні поняття. Легко помітити, що для точного опису об'єктів нам може знадобитися дуже велика кількість вимірювань, не виключено, що і нескінченна. При вивченні об'єктів на основі скінченної кількості вимірювань вектора $x = (x_1, \ldots, x_n)$ може відповідати кілька значень характеристики f. Якщо у дедуктивному підході завжди у підсумку отримуємо істинний вивід, то для обґрунтування індуктивних виводів слід використовувати поняття ймовірності.

Щоб розрізняти такі описи, треба на множині всіх описів об'єктів ввести розподіл ймовірностей. Відомо, що процес вивчення об'єктів шляхом випробувань або вимірювань зручно проводити за допомогою апарату теорії ймовірностей, коли ймовірність подій, що спостерігаються, має частотну інтерпретацію. Надалі для простоти викладу розглянемо булевий випадок, коли кожна ознака набирає двох значень: 0 або 1; тоді є лише 2^{n+1} описів об'єктів, і розподіл ймовірностей на цій множині буде дискретним.

За змістом проблеми індуктивного виводу, розпізнавання і навчання дуже близькі. За допомогою ймовірнісного апарату були формалізовані такі поняття, як клас задач розпізнавання, найкраща функція розпізнавання, процедура розпізнавання та її похибка, множина навчальних вибірок та її ймовірнісні розподіли. Основою індуктивного підходу є відома з курсу теорії ймовірностей формула Байєса.

БАЙЄСІВСЬКІ ПРОЦЕДУРИ
ІНДУКТИВНОГО ВИВОДУ

Поняття незалежності двох або кількох подій (А і В), а також їх умовної ймовірності Р(В|А) посідають, у певному розумінні, центральне місце в теорії ймовірностей. Нагадаємо, що

$$P(B \mid A) = \frac{P(AB)}{P(A)}$$

називають умовною ймовірністю події *B* за умови *A*. З цієї формули маємо

$$P(A \mid B) = \frac{P(A)P(B \mid A)}{P(B)}$$

тобто отримуємо відому формулу Байєса. Байєсівська формула дає змогу оцінювати різні події A_1 і A_2 за їх зв'язком з фіксованою подією *B*. У нашому випадку подія *B* полягає в тому, що вимірювання об'єкта набувають конкретних значень x_1, x_2, …, x_n. Подія A_1 визначається тим, що характеристика *f* набуває значення $f = 0$, а подія A_2 — тим, що $f = 1$. Щоб отримати надалі змістовні результати, потрібно зробити певні припущення щодо ймовірнісних розподілів. Вважатимемо, що розподіл ймовірностей задовольняє умові незалежності ознак x_j для кожного класу об'єктів. Принаймні для медичної експертної системи ця умова виконується. Тоді формула Байєса зв'язує вказані умовні ймовірності певною залежністю.

Байєсівська процедура розпізнавання (або індуктивного виводу) будується на основі того, що кожна з ймовірностей $P(x_j \mid f = i)$ і $P(f = i)$ приблизно замінюється частотою виникнення ознак у навчальній вибірці. Обчислення здійснюються одночасно для всіх стовпців навчальної вибірки.

У байєсівській процедурі характеристика *f* набирає значення того класу, для якого оцінка умовної ймовірності $P(f = i \mid x_1, x_2, \ldots, x_n)$ більша, ніж у іншого.

Вдалося отримати точну оцінку похибки байєсівської процедури розпізнавання. Ця похибка наближається до нуля, тобто достовірність байєсівських процедур розпізнавання зростає із збільшенням розмірів усіх класів навчальної вибірки. Одержаний результат підтверджує факт обґрунтування процедур індуктивного виводу. Доведено також, що байєсівська процедура є оптимальною для всього класу задач розпізнавання.

Формулу Байєса називають ще принципом оберненої ймовірності, або концепцією індуктивної ймовірності. За своєю суттю, вона є аксіомою множення ймовірностей. Таким чином, можна сказати, що велика кількість результатів у теорії ймовірностей є логічним висновком формули Байєса та принципу незалежності подій. Ця формула виражає принцип відносності подій у реальному світі. Тому не дивно, що саме байєсівські процедури виявилися оптимальними у процесі розв'язання задач розпізнавання.

Ключовим моментом обґрунтування процедур індуктивного виводу є структура навчальної вибірки та її ймовірнісні розподіли. Спроби багатьох дослідників побудувати теорію оцінювання процедур індуктивного виводу без урахування цих питань не привели до позитивного результату.

Процедури індуктивного виводу легко узагальнюються на байєсівські мережеві моделі, що, ймовірно, дасть можливість у майбутньому моделювати роботу мозку людини і проектувати навчальні комп'ютери.

У процесі обґрунтування процедур індуктивного виводу вдалося також вирішити таке важливе питання, як інформаційна міра складності задач розпізнавання.

ІНФОРМАЦІЙНА МІРА СКЛАДНОСТІ В ЗАДАЧАХ РОЗПІЗНАВАННЯ

У наш час інформаційна теорія визначається як розділ прикладної математики, який досліджує процеси передачі, зберігання і перетворення інформації. Термін «інформаційна теорія», який виник у 50-х роках минулого століття, не має єдиного загальноприйнятого тлумачення. Основи теорії оптимальної передачі інформації були закладені К. Шенноном. А.М. Колмогоров ввів поняття складності кінцевого об'єкта. Складність він визначав як мінімальну кількість двійкових знаків, які містять усю інформацію про заданий об'єкт, достатню для його відтворення. Він також запропонував три підходи до визначення поняття «кількість інформації»: комбінаторний, ймовірнісний та алгоритмічний.

Щоб отримати змістовні результати у галузі інформаційної теорії, необхідно розуміти, що такі важливі поняття, як інформаційна складність об'єктів та кількість інформації, потрібно розвивати у контексті конкретної проблеми або задачі, в нашому випадку — задачі розпізнавання і навчання.

В задачах розпізнавання інформаційна міра складності є похибкою процедури. Отримано узагальнення байєсівських процедур на схеми із залежними ознаками, які утворюють ланцюг Маркова, а також на ряд інших ймовірнісних розподілів, що дає змогу охопити широке коло практичних застосувань. Таким чином, одержані результати значною мірою розвивають положення А.М. Колмогорова у теорії складності об'єктів.

На відміну від теореми Геделя ми маємо неабиякий позитивний результат: доведено принципову можливість вивчати об'єкти реального світу на основі скінченної кількості вимірювань їхніх характеристик.

ІНДУКТИВНА ЛОГІКА

Відомо, що у дедуктивній математиці логічний вивід будується на основі імплікації «якщо p, то q», яка вважається хибною, коли твердження p є істинним, а q — хибним. Якщо ж p виявляється хибним, то імплікація завжди істинна. Таке поширення не викликає будь-яких труднощів, тому що у процесі логічного виводу ми завжди спираємось на істинні твердження. В математичній логіці прийнято говорити, що з хибного твердження випливає все, що завгодно. Таким чином, логічний вивід обмежений рамками лише істинних тверджень. З огляду на цю обставину в дедуктивному підході неможливо очікувати будь-якого кардинального вдосконалення, і в цьому плані він себе, безумовно, вичерпав.

Якщо характеристика *f* набирає двох значень — істина та неправда, то процедури індуктивного виводу перетворюються на процедури індуктивної логіки. Індуктивна логіка відрізняється від класичної: у ній логічний вивід одержують абсолютно інакше, ніж у дедуктивних системах. Крім того, тут не виконується закон виключеного третього. Це відбувається тому, що один і той же опис об'єкта може одночасно набирати істинного і хибного значення.

Особливість процедур індуктивного виводу полягає в тому, що вони визначаються на основі емпіричної інформації, яка міститься в усіх класах навчальної вибірки. Ще раз підкреслимо: саме це міркування є вирішальним: усі досить сильні дедуктивні системи є неповними, оскільки будуються на основі лише істинних тверджень. Крім того, на відміну від дедуктивних систем, обробка емпіричних даних в індуктивних процедурах здійснюється паралельно для всіх класів навчальної вибірки.

Отримані результати свідчать про універсальність байєсівської процедури індуктивного виводу, яка оптимально переробляє інформацію у навчальній вибірці. Подальший пошук нових оптимальних процедур індуктивного виводу, на наш погляд, має бути пов'язаний з вивченням байєсівських мережевих частково впорядкованих структур, характерних для складних хімічних і білкових сполук.

ПРАКТИЧНЕ ЗАСТОСУВАННЯ ПРОЦЕДУР ІНДУКТИВНОГО ВИВОДУ

У 90-х роках минулого століття завдяки розвитку міжнародних мереж зв'язку і появою сучасних потужних баз даних сталося стрімке відродження експертних систем, призначених для розв'язання різноманітних задач розпізнавання і прогнозування у різних сферах діяльності людини. Експертні системи значною мірою скорочують проведення дорогих експериментів у процесі отримання нових перспективних матеріалів, які використовуються у промисловості й техніці.

Розробка експертної системи, як і інтелектуальне життя людини, складається з двох процесів: накопичення знань та їх використання. Але якщо людина отримує знання з досвіду протягом усього життя, то експертна система на першому етапі одержує свої знання від експертів. У цьому випадку побудова експертної системи є дорогим і тривалим процесом. Коли бракує знань або вони відсутні, процедури індуктивного виводу стають основною компонентою експертної системи.

Досвід розв'язання задач у галузі комп'ютерного матеріалознавства (наплавочних дротів, якісних сталей, матеріалів на основі титану і алюмінію) показав високу достовірність нових методів індуктивного виводу. Вони продемонстрували свою ефективність у процесі розв'язання складних біологічних задач, зокрема таких, як моделювання процесу селекції сільськогосподарських культур і прогнозування білкових сполук.

Отже, крім звичайної, класичної математики, до якої ми звикли ще зі школи, існує інша, на наш погляд, досконаліша — індуктивна математика, яка має серйозні перспективи розвитку. Враховуючи отримані результати, можна з впевненістю сказати, що фундамент сучасної математики та інформаційної теорії треба будувати на нових принципах індуктивної математики, які ми спробували стисло викласти у цій статті.

1. *Рассел Б.* Человеческое познание. — Киев: Ника-Центр, 1997.

2. *Гупал А. М., Пашко С. В., Сергиенко И. В.* Эффективность байесовской процедуры классификации объектов // Кибернетика и системный анализ. — 1995. — № 4.

3. *Вагис Г. А., Гупал А. М., Сергиенко И. В.* Эффективность байесовской процедуры распознавания // Там само. — 2001. — № 1.

4. *Вагис Г. А., Гупал А. М., Сергиенко И. В.* Эффективность байесовской процедуры распознавания. Дискретный случай // Там само. — 2001. — №6.

5. *Гупал А. М., Вагис А. А.* Статистическое оценивание марковской процедуры распознавания // Проблемы управления и информатики. — 2001. — № 2.

6. *Гупал А. М., Сергиенко И. В.* Байесовская процедура — оптимальная процедура распознавания и преобразования информации // Там само. — 2001. — № 3.

7. *Гупал А. М.* К вопросу обоснования процедур индуктивного вывода // Там само. — 2001. — № 6.

І. Сергієнко, А. Гупал

ІНДУКТИВНА МАТЕМАТИКА

Резюме

У зв'язку з неповнотою аксіоматично-дедуктивного підходу показано, що у сучасній математиці рівною мірою повинна розвиватися та використовуватися індуктивна математика. При цьому байєсівський підхід виявляється основою побудови оптимальних процедур індуктивного виводу та індуктивної логіки.

Обґрунтовано, що байєсівські мережі та ланцюги Маркова утворюють джерело нових важливих прикладних застосувань.

I.Sergiyenko, A.Goopal

INDUCTIVE MATHEMATICS

Summary

Due to incompleteness of axiomatic-deductive approach is it shown that inductive mathematics is to be equally developed and applied in modern mathematics.

Bayesian approach is the basis of construction of optimal procedure of inductive inference and inductive logic.

Bayesian networks and Markov shains are the source of new important applications.

© СЕРГІЄНКО Іван Васильович. Академік НАН України. Директор Інституту кібернетики ім. В.М. Глушкова НАН України (Київ).

ГУПАЛ Анатолій Михайлович. Доктор фізико-математичних наук. Директор Науково-навчального центру прикладної інформатики НАН України (Київ). 2002.