

АНАЛІЗ ДАНИХ ЗА УМОВ ТРАНСФОРМАЦІЙНИХ ПЕРЕТВОРЕНЬ – ПРОФЕСІЯ, НАУКА, МИСТЕЦТВО

*Ю.Є. Павленко, к.т.н.
Шосткинський інститут СумДУ*

Процеси розвитку складних систем мають свою специфіку, що суттєво ускладнює їх аналіз. Особливо нелегкими є діагностика стану системи та прогнозування її розвитку. При розробці управлінських рішень реальна дійсність пропонує керівникам вибір – або знайти відповідь на питання: «Чи можна надати незрозумілому факту, феномену більш-менш розумне пояснення?», або намагатися діяти інтуїтивно – на свій страх та ризик.

Отже, разом з реформуванням усіх галузей та сфер суспільного розвитку трансформаційним перетворенням в наш час повинен підлягати і такий вид професійної діяльності як аналіз даних. При цьому під трансформацією потрібно розуміти переорієнтацію методології і технології аналізу та інтерпретації фактів, їх осмислення, наукового передбачення і т. ін. під засади доволі високих стандартів, а також особливості ринкових відносин. Саме тому аналіз даних доцільно вважати не тільки професійною діяльністю, а й наукою та, по великому рахунку, мистецтвом.

Останнє ствердження несе в собі певний дуалізм і означає наступне:

- з одного боку, з метою виявлення у нагромадженні фактів якогось дійсно нового бачення чи тлумачення незрозумілих, «завуальованих» явищ і зв'язків, потрібно задіяти якомога більше наявної інформації;

- з іншого боку, для встановлення головного напрямку розвитку системи, основного закономірного зв'язку, співвідношення необхідно сконцентруватися лише на головних, найсуттєвіших ознаках, безжалісно відкинувши другорядні та несуттєві фактори, з яких складається «прогнозний фон», що перешкоджає виявленню істини – дійсного стану речей.

Саме тому вдосконалення методології аналізу даних є актуальним напрямком прикладних досліджень. Мета роботи – дослідження особливостей та методів аналізу даних, встановлення ступеню їх інформативної спроможності, створення методологічної

бази для розробки новітніх підходів. Методи дослідження – спостереження, аналіз, синтез. Стан будь-якої системи може бути зображений точкою у фазовому просторі, де на осях координат відкладені значення відповідних фазових координат (показників, чинників, умов). Якщо стан системи змінюється у часі, то ця точка переміщується у просторі показників за певною кривою – фазовою траєкторією, або складає фазову площину, фазову поверхню.

Встановлення аналітичної форми зв'язку, тобто вибір та обґрунтування виду рівняння, що описує фазову траєкторію розвитку системи, – найбільш відповідальний етап процесу моделювання соціально-економічних явищ. Математично задача зводиться до створення кореляційної моделі, яка відповідала б певним змістовно-економічним умовам: базувалась на економічній теорії і відображала об'єктивні закономірності та особливості досліджуваних процесів; характеризувала їх зміст і структуру; інтерпретувала економічний зміст як процесу в цілому, так і кожної змінної.

Отже, як на початковій, так і на заключній стадіях етапу специфікації моделі здійснюють змістовний аналіз даних, а тому обов'язково використовують допоміжні, в основному графічні, методи аналізу та якісного тлумачення досліджуваних причинно-наслідкових зв'язків. За допомогою графічних методів забезпечується виявлення та аналіз зв'язків між чинниками (причинами) та характеристиками явища (наслідками) на концептуальному рівні. На ньому дослідник повинен вирішити свою задачу (проблему) у якісному вигляді, тобто, уявити зібрані дані у візуальній формі, „побачити задачу”. Саме „побачити”, оскільки зоровий аналізатор людини – це канал, по якому мозок отримує найбільший обсяг зовнішньої інформації. Власне кажучи, у випадку парного зв'язку цілком можливе впевнене створення характерних проєкцій даних на суттєві ознаки: гістограми розподілу (одновимірної проєкції на ось) та діаграми розподілу (двовимірної проєкції на площину). За наявності відповідних умов вони надають змістовну, а тому корисну інформацію відносно вірогідної форми зв'язку. Крім того, парний зв'язок може інтерпретуватися графічним зображенням регресійних рівнянь. У разі ускладнення процесу досліджуване явище зображують кусочно-лінійною регресією або регресією з точками розриву.

Таким чином, графічне зображення даних є дієвим методом візуальної інтерпретації змістовної сутності складних соціально-

економічних процесів і явищ, а тому його можна вважати, також, засобом опосередкованого впливу на прискорення трансформаційних перетворень у суспільстві.

Отже, парний зв'язок відображається без значних зусиль, а тому у більшості випадків дійсно можливий вибір найбільш придатної форми зв'язку. Що ж стосується множинних зв'язків, то їх графічне зображення потребує неабияких зусиль і є не до кінця вирішеною проблемою.

Простіше за все складається справа з трьохвимірними даними. Так, лінійні зв'язки у трьохвимірному просторі достатньо легко можуть бути представлені у вигляді фазових площин, еквівалентних рівнянням $A \cdot x + B \cdot y + C \cdot z + D = 0$, нелінійні – фазових поверхонь, а мультиплікативні зв'язки, скажімо, виробничі функції типу $Y = b_1 \cdot X_1^{b_2} \cdot X_2^{b_3}$, інтерпретуються (при $b_2 + b_3 = 1$) як фазові конічні поверхні. Проте, в загальній постановці задача формулюється як необхідність інтерпретації множинних зв'язків набором даних, розташованих у якомусь більш складному багатовимірному просторі опису – гіперпросторі.

Яким же чином можливе графічне відображення множинних залежностей, а головне, яка мета повинна переслідуватися при цьому?

Виходячи з того, що гіперпростір – це формальне подання даних, уявляється, що всі об'єкти потрібно перетворити у точки, а всі їх властивості та взаємовідношення відобразити через: координати точок-об'єктів та просторову конфігурацію цих точок у просторі опису вихідних даних. Це означає, що на стадії якісного аналізу необхідно уважно „придивитися” до структури даних, намагаючись якісно, змістовно „побачити задачу”, видобути з зображення неформальний сенс, отримати інформацію про реальні властивості та відношення об'єктів. А для цього потрібно таким чином візуально інтерпретувати первинні дані, щоб вони перетворились у повноцінний матеріал, з якого б структура даних „проглядалась” у явному вигляді і давала відповіді на ряд важливих запитань: яка саме реальність відповідає поняттю „об'єкт дослідження”; які з властивостей є суттєвими; на які пари ознак потрібно спроектувати дані.