

vědeckotechnické revoluce, rozpracovanou v dílech našich a sovětských vědců a nastiňuje reálné cesty zdokonalování všech složek společenského řízení — sociální regulace, jednotlivých druhů řízení i vedení — umožňující zabezpečit komplexní, proporcionální rozvoj společnosti a cílevědomě využívat výsledků vědy i moderní techniky pro šťastný a bohatý život lidu. Škoda, že do této části práce autor v konkrétní podobě nezakomponoval poznatky, které přinesla kategoriální analýza řízení a které mohou dále přispět k zdokonalení jednotlivých složek řídicího procesu — plánování, programování, projektování.

M. Pittner dále podrobuje kritice buržoazní názory na řízení vědy, zejména jednostrannost internalistického a externalistického pojetí a dokazuje, že efektivně řídit vědu v zájmu celé společnosti i pracujícího člověka dokáže jedině socialismus. Je tomu tak proto, že za socialismu společenské řízení cílevědomě usměrňuje rozvoj vědy tak, že sjednocuje „vnější“ společenské působení na vědu, která dostává charakter vědecké orientace na základě společenských potřeb, s vývojem „zevnitř“, s vlastní orientací vědců, jež nabývá podoby jejich společenské sebereflexe, společenského sebeuvědomění a společenského sebeřízení. Čím důsledněji řízení zabezpečuje, aby věda plnila své rostoucí společenské funkce, aby co nejlépe přispívala k rozvoji výrobních sil, společenských vztahů i ke všeobecnému rozvoji člověka, tím více se v socialismu utvářejí příznivé společenské podmínky pro efektivní řízení jejího vlastního rozvoje, pro realizaci dalekosáhlých změn v její povaze, spojených s pronikáním do nových sfér a vrstev sociální skutečnosti (s. 196).

Ke kladům monografie patří, že autor přistupuje k výzkumu řízení netradičním způsobem, že nastoluje nové otázky a snaží se na ně kvalifikovaně odpovědět. Nejvýraznější se to projevuje v podrobné analýze základních kategorií — regulace, řízení, vedení — jejich obsahu a vzájemných dialektických vztahů. M. Pittner tvořivě používá marxistický filozofický aparát při výzkumu řízení jako sociálního fenoménu i při rozpracovávání některých závažných otázek — například vztahu mezi subjektem a objektem řízení, úlohy racionality, účelu a cíle v řízení, vztahu mezi živelností a řízením, charakteru vztahů v systému řízení i dalších — jejichž řešení umožňuje prohlubovat světónázorovou i teoreticko-metodologickou stránku marxisticko-leninské vědy o řízení. Podnětné jsou z tohoto hlediska jeho náměty na rozvíjení jednotné vědy o řízení socialistické společnosti i nástin místa a úlohy jednotlivých, zejména společenských věd v tomto procesu. Kritická analýza některých buržoazních teorií řízení umožnila hlouběji osvětlit jejich ideologické a apologetické poslání. Odhalení odcizeného charakteru buržoazního řízení patří k teoretickým přínosům práce.

Positivním rysem je, že autor se snaží uplatnit kategorie, které rozpracovává z obecné teoretického hlediska při řešení některých konkrétních problémů — při analýze vztahů mezi politikou, řízením a vedením v buržoazní a socialistické společnosti, mezi vědou, technikou a řízením. Podařilo se mu přesvědčivě ukázat, že při komplexním přístupu k celospolečenskému systému řízení nelze opomíjet úlohu regulačních procesů v celém systému i v jeho jednotlivých složkách, že odpovídající místo je nutno věnovat funkcím různých druhů řízení, jakož i způsobům a formám vedení, chceme-li vytvořit adekvátní systém řízení, schopný zabezpečit harmonický a proporcionální vývoj socialistické společnosti.

Avšak otázka, jak uplatnit obecné poznatky při zkvalitňování řízení v jednotlivých společenských oblastech i různých sférách společenského řízení, byla zodpovězena jen částečně. Její podrobné řešení bude vyžadovat hlubší analýzu. Nutno však ocenit autorův přístup k velice aktuální společenské problematice, který umožnil z marxistických pozic obohatit výzkum řízení o nové momenty i pohledy, jež přispívají k rozvoji vědy o řízení socialistické společnosti.

Jiří Kuzin

L. Lebart, A. Morineau, K. M. Warwick: Multivariate Descriptive Statistical Analysis
(Mnohorozměrná deskriptivní statistická analýza)

New York, John Wiley & Sons 1984, 231 s.

Jak patrně z podtitulu *Korespondenční analýza a příbuzné techniky pro velké matice*, kniha se zabývá jednou z nejprogresivnějších a nejnadanějších metod a představuje autoritativní a autentický záznam o stavu jejího rozvoje a praktické využitelnosti.

Deskriptivní metody ve statistické analýze dat nabývají na důležitosti především ze dvou důvodů: 1) věda, výzkum i praxe mají k dispozici velké množství údajů, které je třeba přehledně uspořádat, zredukovat a zhodnotit; 2) vznik datových souborů se většinou neřídí podle modelů pravděpodobnostních výběrových plánů, stále častěji se objevují data úplných šetření, u nichž je výběrová chyba nepatrná, zato však jiné typy chyb výrazné.

Výhodnost deskriptivních metod spočívá v tom, že nejsou zatížené nereálnými předpoklady (jako je např. mnohorozměrná normalita) a že odpovídají na většinu úloh společenskovedního výzkumu, v němž se stále více prosazuje metodologie explorační analýzy. Ptáme se, jaká je struktura dat, jaké jsou mezi nimi vztahy, jaké pravidelnosti, abychom následně tyto datové jevy zhodnotili, sociologicky rozebrali a interpretovali, formulovali hypotézy a závěry. Opačný případ, kdy předem formulujeme sociologické hypotézy postihující strukturu dat, jež má být ověřena či zamítnuta, je relativně vzácný.

V analýze dat společenskovedních výzkumů stále dominuje typ velkého souboru. 15–20 tisíc jedinců není výjimkou, výzkumy veřejného mínění mají běžně 2000–3000 respondentů, počet proměnných bývá 200, ale také 500 a po transformacích i více. Autoři recenzované knihy popisují právě takové postupy, které umožňují explorační analýzy velkých datových masivů, čímž se myslí mnoho jednotek, proměnných i kategorií. Metody mnohorozměrné deskriptivní statistické analýzy (MDSA) vycházejí z možností, které poskytl moderní počítač a pod jejich vlivem vyvíjená metodika numerického zpracování.

Cílem procedur MDSA je určení vztahů mezi proměnnými, mezi kategoriemi proměnných a mezi jednotkami v přehledné a jednoduché formě (číselné i grafické). K tomu se využívá jednak dělení souboru na části, v nichž jsou si jednotky vzájemně velmi podobné a proměnné působí shodně, a jednak grafického umístění jednotek, proměnných a kategorií v rovině (rovinách) tak, že každá z nich je reprezentována bodem. Přitom blízkost u proměnných a kategorií označuje shodnou roli a kladnou korelovanost, zatímco velká vzdálenost vyjadřuje opačné působení, zápornou korelovanost. Umístění bodů jednotek umožňuje rychle rozpoznat jednotky podobné a nepodobné.

Obsah knihy je rozvržen do osmi kapitol, v nichž jsou popsány metody, vyvinuta teorie, uvedeny praktické ilustrace, rozebrány metodologické a obsahové otázky a nakonec přetištěn výpis výpočetního programu s krátkým uživatelským návodem.

Kapitola I, Popisná analýza hlavních komponent a kanonické rozklady (Descriptive Principal Component Analysis and Singular Value Decomposition), předkládá metodiku zpracování vektorů spojitých číselných veličin. Metoda hlavních komponent může být chápána jako jedna z technik faktorové analýzy. H. Hotelling ji zavedl k vyhledání takové funkce proměnných, která na sebe váže maximálně možné množství informace z celého datového souboru v tom smyslu, že nejlépe rozlišuje mezi jednotkami na jednorozměrné škále. Postupně lze nalézt několik takových nezávislých a sekvenčně optimálních komponent.

V souvislosti s tímto základem autoři úlohu takto formulují: nalézt jasnou vizuální (geometrickou i číselnou) reprezentaci matice dat tak, aby mohl výzkumník snadno a rychle identifikovat příbuznosti mezi jednotkami, vztahy mezi proměnnými a odhalit strukturu datového souboru.

Matematická podstata problému je diskutována podrobně, neboť je základem i pro řešení modelů dalších kapitol. Je to hledání charakteristických čísel a vektorů symetrických pozitivně semidefinitních matic, tzv. *diagonalizace matice*. V kapitole se též rozebírá problém standardizace proměnných a použití Spearmanových pořadových koeficientů korelace pro neparametrickou analýzu.

Obbohacím analýzu dat je zavedení *dodatečných proměnných*, jednotek a kategorií (nazývají se také *doplňkové, ilustrační*), které se neúčastní základních odvození hlavních komponent. *Dodatečné jednotky* jsou pomocí již známých skórovacích vzorců umístěny mezi původní (aktivní nebo základní) jednotky, s nimiž tak mohou být jak číselně, tak graficky porovnávány. Umístění dodatečných proměnných v konfiguraci původních aktivních proměnných ukazuje na korelovanost a vztahy s odvozenými hlavními komponentami.

Kapitola II, Korespondenční analýza (Correspondence Analysis), řeší obdobnou úlohu jako předcházející, avšak roli proměnných zde hrají kategorie jedné proměnné. Je zde popsán model, který vysvětluje vztahy v dvourozměrné kontingenční tabulce. Analýza je založena na chi-kvadrátové vzdálenosti podmíněných rozdělení řádků a sloupců.

Kapitola III, Kanonická a diskriminační analýza – teoretické a technické úvahy: stručný přehled (Canonical Analysis and Discriminant Analysis – Theoretical and Technical Considerations: A Brief Review), uvádí model diskriminační analýzy, kanonické korelační analýzy (a vícenásobnou regresi jako její speciální případ), jejich vzájemnou modelovou vazbu i vztah ke korespondenční analýze. Kapitola má nejen značný význam teoretický, ale její výsledky ukazují, jak lze využít programy pro kanonické korelace a diskriminační analýzu na úlohy korespondenční analýzy, když nejsou speciální programy k dispozici.

Kapitola IV, Mnohonásobná korespondenční analýza (Multiple Correspondence Analysis), patří k nejdůležitějším a tvoří vlastně jádro knihy. Popisovaná metoda je velkým přínosem praxe, neboť zajišťuje souběžné vyhodnocení velkého počtu kategorizovaných proměnných s mnoha kategoriemi na rozsáhlém souboru jednotek. Výsledkem je popis dat, obdobně jako u hlavních komponent, určení postavení jednotek ve struktuře souboru

a odhalení relevance jednotlivých kategorií vzhledem k jejich vlivu na heterogenitu souboru. V kapitole je uváděn vztah mezi analýzou matice dat a Burtovy tabulky (plošné uspořádání kontingenčních tabulek několika proměnných) i kanonickými korelacemi. Řešení úlohy je tak podáno v několika variantách. Najdeme tu i přibližné testy pro významnost vlivu dodatkových (ilustračních) kategorií.

Kapitola V, Automatická klasifikace – seskupovací techniky užité s metodou hlavních os (Automatic Classification – Clustering Techniques Used with Principal Axes Method), uvádí strategii seskupování, kterou autoři vyzkoušeli a o níž zdůrazňují, že není teoreticky optimální, ale velmi výhodná prakticky, protože ji lze použít na velké soubory. Nejde tedy o přehled seskupovacích metod, ale o jednu prověřenou a doporučovanou metodu, která sestává z kombinace několika přístupů. Její postup lze vyjádřit v krocích:

1. Rozklad souboru na stabilní skupiny.

- a) Rozklad souboru na zvolený počet tříd, vycházející z náhodných středů a přiřazující jednotky k těmto středům podle minimální vzdálenosti. Iterační zlepšení rozkladu k lokálnímu minimu.
- b) Několikanásobné opakování rozkladu z jiných náhodných středů.
- c) Zjemnění všech výsledných rozkladů. Třídy zjemněného rozkladu jsou nazvány stabilní skupiny.

2. Sloučení stabilních skupin pomocí Wardova kritéria.

3. Optické určení počtu výsledných skupin řezem v dendrogramu.

Spolu s popisem algoritmů tu nalezne čtenář i doporučení na volbu parametrů, které plyne z dlouhého experimentálního ověřování. Pozornost je věnována popisu výsledných skupení a grafickému zobrazení jejich vztahu k výsledkům korespondenční analýzy, neboť postup se používá (v kontextu knihy) na údaje odvozené z metody hlavních komponent nebo korespondenční analýzy (faktorová skóre).

Kapitola VI, Algoritmy při přímém čtení (Direct Reading Algorithms), věnuje pozornost algoritmické stránce metod, kterou se autoři zabývají velmi pečlivě a objevně. Uživatele-sociologa tyto problémy přímo nezajímají, ale nepřímě je na nich závislý. Algoritmizace musí nechat informaci nezkrácenou, dovolit analýzu velkých souborů při nízkých nákladech a malé spotřebě strojového času a umožnit používání dostupných počítačů. Tato důležitá činnost matematiků je před sociologem většinou skryta a je jím též obvykle podceňována. Sociolog vidí až výsledek a cesta k němu je mimo jeho zájem. A přesto tato nepřímá závislost by měla vést k vyšší podpoře i této čistě technické stránky věci.

Kapitola VII, Spolehlivost a významnost výsledků (Reliability and Significance of Results), odpovídá na otázky: Jaké matice dat analyzovat pomocí MDSA? Jak je získávat? Co můžeme očekávat od technik MDSA? Jak hodnotit kvalitu získaných konfigurací? Například odpověď na první (základní) otázku zní: 1. tak velké, že vizuálně nebo elementární statistickou analýzou jejich strukturu nelze odhalit; 2. obsahově homogenní, takže má smysl počítat vzdálenosti mezi řádky a mezi sloupci a vzdálenosti jsou smysluplně interpretovatelné; 3. apriori amorfní, tj. takové, u nichž struktura není předem známá.

Kapitola VII, Počítačový program: korespondenční analýza pro velké matice (A Computer Program: Correspondence Analysis for Large Matrices), umožňuje čtenáři provádět vlastní výpočty vícenásobnou korespondenční analýzou. Obsahuje výpis programu CORAN (22 subrutin v jazyce FORTRAN IV, 1386 štítků včetně komentářů), popis programu, popis příkazů, odhadové vzorce pro potřebnou paměť počítače, zkušební příklad s výstupy. Jsou tu také popsána možná zjednodušení. CORAN obsahuje dvě etapy z širšího programu SPAD, který vyvinuli L. Lebart a A. Morineau se svými spolupracovníky. Je to ovšem samostatný program, který poskytuje i nutné numerické a grafické procedury. Může porovnávat binární datové matice (kódování kategorií dat jako 0, 1, jeden sloupec je jedna kategorie), výpočty lze doplnit zkrácenými názvy identifikujícími řádky i sloupce (jednotky, proměnné i kategorie), zpracovává aktivní i doplňkové proměnné. Při velkém počtu kategorií však CORAN nemůže konkurovat s efektivitou programu SPAD, který čte přímo redukovaná kódování číslem kategorie u každé proměnné, jak je u nás běžným zvykem (jeden sloupec je jedna proměnná, kód je číslo kategorie).

Korespondenční analýza se pod různými názvy vyskytuje u mnoha autorů, vznikají různé směry a celé školy. L. Lebart a A. Morineau jsou v této oblasti známými autory, jejichž práce již dnes značně přesahují rámec, v němž se pohybuje obsah knihy. Jejich přístup je založen především na geometrické představě projekce v euklidovském prostoru s cílem „poskytnout geometrickou reprezentaci informace“. Dalším zdůrazněným aspektem metodiky je možnost práce s neagregovanými daty, tedy analýza původních vstupních jednotek, určení jejich vztahu k ostatním jednotkám i pozice v celkové struktuře souboru. S tím souvisí vysoká pozornost přípravě i programové realizaci numerických aspektů metodiky. Ovšem i analýzy tabulek četností jsou pojednány dobře a dostatečně.

To vše patří k nesporným kladům knihy samé i francouzského směru v korespondenční analýze, který reprezentuje.

Jiný přístup se stejným matematickým základem kanonických rozkladů spočívá ve snaze o optimální souběžnou kvantifikaci řádků a sloupců tabulky, což je charakteristické pro japonskou školu, vzniklou kolem prací Hyashiho. Tomu odpovídá také název „duální škálování“, propagovaný S. Nishisatem (viz recenzi jeho knihy v Sociologickém časopise 1985, č. 2, s. 212). Rozdíl v přístupech obou monografií je v tom, že Nishisato se důsledně zabývá kategorizovanými daty v různých formách tabulek v kontextu psychologických dat v učebnicovém podání, zatímco Lebart, Morineau a Warwick uvádějí korespondenční analýzu v širším pojetí a s přímou návazností na „hrubou“ analytickou práci.

U nás od poloviny sedmdesátých let publikují výsledky o metodice příbuzné korespondenční analýze J. Řehák a I. Loučková. Jejich přístup (viz např. Sociologický časopis 1983, s. 535; 1984, s. 174; 1985, s. 176) se neomezuje jen na kategorizovaná data, ale směřuje k analýze libovolných, ale obsahově homogenních obdélníkových tabulek dat nebo statistik. Nehledají však příbuznost mezi maticí dat a z ní odvozenými tabulkami. Jejich přístup umožňuje souběžnou interpretaci faktorovou (a tudíž případně i kauzální), interpretaci z hlediska modelu mnohorozměrného škálování a kvantifikační (model reciprokých průměrů). U kontingenčních tabulek nevycházejí ze vzdálenosti chí-kvadrátového typu, ale doporučují dvojité centrované matice logaritmovaných četností. Na rozdíl od jiných přístupů věnují též pozornost různým typům rotace.

Knihy Lebart, Morineau a Warwicka je velmi vítaná, neboť rozšiřuje analytické obzory a poskytuje návod na řešení úloh, které začínají ve společenských vědách dominovat. Užitečnost aplikací korespondenční analýzy prokázala již řada publikovaných prací (i u nás). Výzkumníci by se neměli korespondenční analýzy zříkat, přestože jde o metodu, jejíž ovládnutí vyžaduje studium a praxi, jejíž výsledky mohou být někdy i nesnadno interpretovatelné a na niž nejsme zatím ještě zvyklí. Odhalená struktura může být někdy i značně překvapivá nebo jinými metodami neidentifikovatelná.

Recenzovaná kniha reprezentuje velmi silný a nadějný trend ve vývoji statistické metodologie a bylo by velmi žádoucí, kdyby se naši výzkumníci i statistikové mohli s postupy seznámit podrobně a metodu v praxi používat. V NDR přeložili podobně zaměřenou knížku, na níž se spolu s Lebartem a Morineauem podílel I. P. Fenelon (*Traitement des Données Statistiques. Méthodes et Programmes*, Dunod, Paris, 1979). Byla velmi rychle rozebraná a vzbudila značný zájem metodologů i výzkumníků. I u nás by měla podobná práce vyjít, abychom udrželi krok se světovým vývojem v oboru analýzy dat.

Blanka Řeháková