

Dvě knihy /o základech mnohorozměrného škálování

A. P. M. Coxon: The User's Guide to Multidimensional Scaling with Special Reference to MDS(X) (Uživatelský průvodce mnohorozměrným škálováním se zvláštním odkazem na MDS(X))

London, Heinemann Educational Books 1982, XII + 271 s.

P. M. Davies — A. P. M. Coxon (eds): Key Texts in Multidimensional Scaling (Klíčové texty v mnohorozměrném škálování)

London, Heinemann Educational Books 1982, XX + 349 s.

Mnohorozměrné škálování patří již dlouhou dobu neodmyslitelně k metodologii a metodice analýzy dat ve společenských vědách. U nás je používáno v sociologii a ve výzkumu veřejného mínění po více než deset let. Z bohatých možností výběru modelu a typů vstupních dat i výstupních forem je však využívána jen nepatrná část. Velké bohatství možností nabízí programový systém MDS(X), o němž pojednala recenze J. Pavlíka (Sociologický časopis 1987, č. 3, s. 303). Recenze obsahovala též popis modelů, které jsou pomocí tohoto systému aplikovatelné.

Experimentování s modely MDS(X) potvrzuje, že metodika zasluhuje uplatnění v plné šíři. Její závažnost pro praxi poznávacího procesu, její analytické možnosti i její přínos k formám výstupů jsou indikovány nejen stálým rozvojem a neustále se objevujícími zdokonaleními, modifikacemi, novými modely, ale i množstvím aplikací v odborné literatuře světové vědy.

Obě recenzované knihy kladou důraz na *nemetrické* metody, tj. takové postupy, které je možno také charakterizovat jako neparametrickou (pořadovou) analogii u nás známých a používaných metod klasického mnohorozměrného škálování (DISTAN), faktorové analýzy, metody hlavních komponent, korespondenční analýzy, metody LINDA apod. „Nemetrické“ znamená, že číselné hodnoty vstupů jsou brány jako relativní, resp. pořadové, že se s nimi nepracuje absolutně ve významu číselných vzdáleností s vlastnostmi metriky. Tak na vstupu mohou být použity různé koeficienty asociace, korelace či podobnosti, které přímo do klasické vícerozměrné statistické analýzy vstoupit nemohou, nebo koeficienty, které sice metrické vlastnosti mají, ale výzkumník nechce v daném případě riskovat nedostatky přístupu nejmenších čtverců či dalších klasických předpokladů. Smyslem mnohorozměrného škálování je (a) metrize, (b) grafická podoba výsledků, (c) redukce informace, (d) extrakce latentních proměnných (faktorů).

Přínos obou knih je nesporný. Uživatelská příručka A. P. M. Coxona ukazuje přehledně na stav teorie, metodologie a postupů, který odpovídá začátku a první polovině osmdesátých let v tomto oboru. Je to kompendium pojmů a metod. Uživateli je postupně veden dnes již obsáhlým a poněkud nepřehledným, v časopisecké literatuře existujícím labyrintem postupů a modelů. Autor systematicky, krok po kroku, zdánlivý chaos metodologie mnohorozměrného škálování rozptýluje. Tak dostává čtenář jednak přehled o principech a modelech, jednak je veden ke správnému výběru metody — a tím ovšem také ke správné interpretaci výsledných výstupů. U nás byla, právě pro rozptýlenost, nejednotnost a nepřehlednost časopiseckých publikací, aplikována de facto pouze metoda SSA (Smallest Space Analysis) převádějící libovolné míry podobnosti (nepodobnosti) do metrických vztahů eukleidovské roviny nebo prostoru. Coxonova kniha je vhodným úvodem, který může k překonání této bariéry přispět. Je plná názorných schémat, tabulek, přehledů, obrázků, které slouží jako shrnutí, ilustrace, ale také jako kroky v rozhodování o volbě metody. To samo je dosti k tomu, aby se monografie zařadila mezi úspěšné metodologické tituly. Tím spíše, že jako její paralela existuje užitečný a souhrnný systém programů MDS(X), který nabídku knihy realizuje (viz výše uvedenou recenzi J. Pavlíka). Nejde tedy pouze o abstraktní látku, ale o praktické možnosti. Obsah je tvořen osmi kapitolami:

1. Úvod do mnohorozměrného škálování
2. Data: míry podobnosti a nepodobnosti
3. Škálování: základní nemetrický distanční model
4. Interpretace konfigurací
5. Charakteristiky MDS modelů
6. Programy pro analýzu dvourozměrných tabulek dat
7. Třírozměrné a další rozšíření základního modelu
8. Vývoj v MDS

Knihy není přehlcena technickými detaily, ale podává plnou informaci o matematických principech modelů. Algoritmy jsou popsány verbálně, pomocí schémat

výpočetních kroků a pomocí příkladů, které na ukázkách předvádějí iterační změny. Kniha tak slouží zároveň jako úvod do teorie mnohorozměrného škálování i pro vážnější zájemce o matematickou stránku. Dalším kladným rysem této monografie jsou vhodně vybrané příklady z publikované odborné literatury (některé z vlastní vědecké práce autora), které jsou podrobně rozebírány.

Mnohorozměrné škálování vychází z psychometrických teorií a dodnes je jeho terminologie zatížena terminologií vzniku a tím, že jeho vývoj probíhá především na stránkách psychometrických periodik. Je také různá u různých metod. Autor se této tradice drží, přebírá terminologii v konzistenci s ostatní literaturou, sjednocuje ji (i když ne s plným úspěchem) a i tím přispívá k odstraňování jedné z bariér porozumění u širokého okruhu zájemců o metodu. Psychometrická terminologie však sama o sobě takovou bariérou je (mezi uživateli mimo psychologii). Formulace v termínech stimulů, subjektů, skupin apod. věc zatemňuje spíše než ji pomáhá (na rozdíl od autorova názoru a záměru, jak vyplývá z poznámky na s. 190).

Zvládnutí jednotlivých metod může být nespecialistům ulehčena především ne-neutrální a sjednocenou terminologií (nejlépe matematické statistiky a jazyka velkých programových systémů). Vždyť aplikace v sociologii, geografii, biologii, ekonomii nemají s původním psychologickým modelem nic společného (jde o období vývoje faktorové analýzy). Jsem přesvědčen, že rozvoj metodologie mnohorozměrného škálování dosáhl takového stupně, že je třeba k takovému kroku již přistoupit a uživatelům maximálně zjednodušit situaci. Alespoň v české a slovenské terminologii bychom se měli přímému převodu psychometrických termínů vyhnout.

Druhý svazek je sborníkem statí, redigovaný P. M. Daviesem a A. P. M. Coxem. Je to „čítanka“ (reader) v metodách mnohorozměrného škálování, sborník sedmnácti závažných, dnes většinou již klasických článků, v nichž myšlenky, modely a algoritmy vznikaly, byly rozvíjeny a aplikovány. Výběr a kompozice takových čítanek v oboru je vždy subjektivní, je vždy obtížné vybrat 15–20 statí ze sbírek, které byly publikovány. Metodologie škálování se rozvíjí přibližně 20–25 let, za nejdůležitější momenty lze považovat asi práce P. N. Sheparda z roku 1962 (*Psychometrika* 27, s. 125–139, 219–248), myšlenkově bohatou monografii C. Coombse: *Theory of Data* (N. Y., J. Wiley) z roku 1964 a souběžné práce J. B. Kruskala v téže roce (*Psychometrika* 29, s. 1–27, 115–129; obě jsou ve sborníku přetištěny). Od tohoto přelomu se objevilo velké množství teoretických, metodologických i aplikačních prací. Proto sestavení sborníku není snadnou záležitostí; vyžaduje hlubokou i širokou znalost rozptýlené literatury, celkových výsledků oboru v jeho různých vývojových směrech a současné trendy v nich. Zde se podařilo sestavit velmi užitečný sborník obsahující to nejdůležitější a nejzákladnější. Jsou tu zařazeny přehledové články i práce uvádějící ve své době originální a nové prvky vyvíjené teorie. Jsou to práce o metodě i interpretaci.

Sborník je rozdělen na pět částí:

- I. Orientace
- II. Základní model
- III. Interpretace
- IV. Třetí rozměr
- V. Nový vývoj

Obě knihy tvoří jeden celek, ač každá má svůj vlastní samostatný význam. Výtečně se doplňují, klíčové texty jsou doplňkem monografie *Uživatelský průvodce*. Třetí součást celku je programový systém MDS(X) se svým manuálem.

Knihy shrnují základy, a proto jejich platnost a užitečnost pro čtenáře v aplikačních oborech není nijak časově omezena. Postupně budou překonány části *Nový vývoj*, které jsou ovšem časově podmíněny. Oba svazky mohou doporučit ke studiu. A jelikož je načase, aby noví adepti sociologické profese dostávali metodologické vzdělání podle vývoje oboru ve světě, bylo by vhodné, aby obě knihy byly využity ve výuce.

Poznamenejme, že programy mnohorozměrného škálování se stávají součástí statistických systémů pro personální počítače již ve velmi brzkém stupni jejich vývoje, dříve než metody seskupování, log-lineárních modelů, mnohorozměrné analýzy rozptylu apod. To nejen ukazuje na požadavky ze strany kupujících (programy uvedeného typu vznikají na komerčním základě, a proto pořadí při jejich vybavování není v žádném případě nahodilé), ale také zpětně dále přispívá k rychlému rozšíření. Metodika mnohorozměrného škálování ve velkých systémech pro společenskou vědu je již dnes samozřejmostí (viz procedura ALTAB v SPSSX). I pro sociologii platí, že se musí dostat na úroveň světové vědy — k tomu musí používat nejvyšších prostředků a nejnovější metodologie.

Jan Řehák