Kybernetika

Jiří Skolka

Použití míry množství informace při agregaci bilancí meziodvětvových vztahů

Kybernetika, Vol. 1 (1965), No. 1, (62)--73

Persistent URL: http://dml.cz/dmlcz/124840

KYBERNETIKA ČÍSLO 1, ROČNÍK 1/1965

Použití míry množství informace při agregaci bilancí meziodvětvových vztahů

JIŘÍ SKOLKA

V článku se ukazuje možnost použití Shannonovy míry entropie k řešení ekonomického problému — agregace meziodvětvových bilancí. Z ekonomického hlediska dává použitá metoda uspokojivé výsledky.

MATEMATICKÝ MODEL MEZIODVĚTVOVÝCH VZTAHŮ

Bilance meziodvětvových vztahů

Bilance meziodvětvových vztahů je zvláštním typem národohospodářské bilance. Hlouběji než ostatní vrcholné bilance zobrazuje strukturu tvorby a rozdělování společenského produktu a národního důchodu. Navíc lze vztahy mezi jejími položkami matematicky vyjádřit a analyzovat. Proto se bilancování meziodvětvových vztahů, jehož hlavní zásady formuloval koncem třicátých let americký ekonom ruského původu W. Leontief, rychle rozšířilo a je dnes běžnou činností statistických úřadů většiny vyspělých zemí.

V této stati pojednáme o jednom problému meziodvětvové analýzy, o problému vhodné agregace odvětví v bilanci, a ukážeme, jak lze k jeho řešení využít elementárních poznatků teorie informace.

Obsah meziodvětvové bilance a vztahy mezi jejími položkami nejsnáze pochopíme v obecném algebraickém vyjádření. Zavedeme tyto pojmy a jim odpovídající symboly:

n počet odvětví materiální výroby, na která bylo národní hospodářství v bilanci rozděleno;

m počet složek přidané hodnoty (Přidanou hodnotou rozumíme rozdíl mezi celkovou hodnotou výroby odvětví a materiálovými náklady. Představuje součet amortizace, mezd, zisku resp. ztrát.);

x_i hodnota celkové produkce (tj. hodnota veškeré výroby bez ohledu na její další použití) i-tého odvětví materiální výroby ($i = 1, 2, \dots, n$);

y_i hodnota konečné produkce (tj. součet spotřeby obyvatelstva, investic a salda dovozu a vývozu) i-tého odvětví materiální výroby;

x_{ij} výrobní spotřeba výrobků i-tého odvětví v j-tém odvětví ($i, j = 1, 2, \dots, n$);

z_k velikost k-té složky přidané hodnoty ($k = 1, 2, \dots, m$);

z_{kj} velikost k-té složky přidané hodnoty, vynaložené v j-tém odvětví;

x součet celkové produkce všech odvětví (společenský produkt).

Vztahy v řádcích a sloupcích meziodvětvové bilance popisují dvě soustavy rovnic. První popisují rozdělení celkové produkce jednotlivých odvětví a nazývají se distribučními rovnicemi:

$$
\begin{array}{c}
x_1 - x_{11} - x_{12} - \dots - x_{1n} = y_1 \,, \\
x_2 - x_{21} - x_{22} - \dots - x_{2n} = y_2 \,, \\
\vdots \quad \vdots \quad \vdots \qquad\quad \vdots \quad \vdots \\
x_n - x_{n1} - x_{n2} - \dots - x_{nn} = y_n \,, \\
\hline
z_1 - z_{11} - z_{12} - \dots - z_{1n} = 0 \,, \\
\vdots \quad \vdots \quad \vdots \qquad\quad \vdots \quad \vdots \\
z_m - z_{m1} - z_{m2} - \dots - z_{mn} = 0 \,.
\end{array}
\tag{1}
$$

Druhé popisují strukturu nákladů ve sloupcích a nazývají se rovnicemi nákladovými:

$$
\begin{array}{c}
x_{11} + x_{21} + \dots + x_{n1} + z_{11} + \dots + z_{m1} = x_1 \,, \\
x_{12} + x_{22} + \dots + x_{n2} + z_{12} + \dots + z_{m2} = x_2 \,, \\
\vdots \quad \vdots \qquad \vdots \quad \vdots \qquad \vdots \quad \vdots \\
x_{1n} + x_{2n} + \dots + x_{nn} + z_{1n} + \dots + z_{mn} = x_n \,.
\end{array}
\tag{2}
$$

Na první soustavě lze ukázat dělení meziodvětvové bilance na čtyři základní části, zvané kvandranty. První kvadrant obsahuje proměnné x_{ij} a popisuje strukturu výrobní spotřeby, druhý kvadrant obsahuje proměnné y_i a popisuje strukturu konečné produkce (může obsahovat i několik sloupců, je-li konečná produkce členěna podrobněji podle způsobu použití). Třetí kvadrant popisuje strukturu přidané hodnoty a je představován proměnnými z_{kj}. Čtvrtý kvadrant zobrazuje znovurozdělení národního důchodu (např. použití daní placených obyvatelstvem, použití zisků podniků apod.). V praxi se zaplňuje jen zřídka, zpravidla chybí evidenční podklady. Soustava rovnic (1) zobrazuje takto zjednodušenou bilanci a obsahuje ve čtvrtém kvadrantu nulové prvky.

Soustavy rovnic (1) a (2) popisují, ale nevysvětlují vztahy mezi položkami meziodvětvové bilance. Chceme-li např. zjistit, jak změna jedné položky bilance zapůsobí na položky ostatní, nemůžeme pracovat jen s jejími údaji, ale musíme použít matematického modelu meziodvětvových vztahů.

Matematický model meziodvětvových vztahů vychází z poznatku, že z různých závislostí mezi položkami meziodvětvové bilance jsou nejvýznamnější a nejstabilnější relace mezi objemem celkové produkce a náklady na produkci. Roste-li nebo klesá-li objem produkce, jednotlivé nákladové položky se úměrně zvyšují či snižují. V praxi tento vztah zpravidla platí a umožňuje definovat ukazatele, nazývané koeficienty přímých nákladů (někdy též technické koeficienty):

(3a)
$$a_{ij} = \frac{x_{ij}}{x_j},$$

(3b)
$$a'_{kj} = \frac{z_{kj}}{x_j}.$$

Dosazením koeficientů přímých nákladů do soustavy rovnic (1) dostáváme soustavu nehomogenních lineárních rovnic, která obsahuje $n(n + m)$ koeficientů a_{ij} a a'_{kj} (o kterých předpokládáme, že jsou stálé), n proměnných x_i, n proměnných y_i, a m proměnných z_k, celkem tedy $2n + m$ proměnných. Koeficienty této soustavy rovnic tvoří prvky dvou matic, jedné čtvercové typu $n\,.\,n$, druhé obdélníkové typu $m\,.\,n$:

(4a)
$$\mathbf{A} = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix},$$

(4b)
$$\mathbf{A}_z = [a'_{kj}] = \begin{bmatrix} a_{11} & a'_{12} & \dots & a'_{1n} \\ \vdots & \vdots & & \vdots \\ a'_{m1} & a'_{m2} & \dots & a'_{mn} \end{bmatrix}.$$

Údaje o celkové a konečné produkci odvětví a o celkové velikosti složek přidané hodnoty tvoří dva n-složkové a jeden m-složkový sloupcový vektor:

(5a, b, c)
$$\mathbf{x} \equiv \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_j \\ \vdots \\ x_n \end{bmatrix}, \qquad \mathbf{y} \equiv \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_i \\ \vdots \\ y_n \end{bmatrix}, \qquad \mathbf{z} \equiv \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_k \\ \vdots \\ z_m \end{bmatrix}.$$

Matematický model meziodvětvových vztahů pak tvoří dvě maticové rovnice:

(6a) $$(\mathbf{E} - \mathbf{A})\,\mathbf{x} = \mathbf{y}\,,$$

(6b) $$\mathbf{A}_z\mathbf{x} \qquad = \mathbf{z}\,,$$

které lze souhrnně psát:

(6c) $$\begin{bmatrix} \mathbf{E} - \mathbf{A} \\ \\ \mathbf{A}_z \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{y} \\ \\ \mathbf{z} \end{bmatrix}$$

($\mathbf{E}$ je jednotkovou maticí typu $n \,.\, n$).

Matematickým modelem meziodvětvových vztahů lze řešit dva základní problémy bilancování výroby. Je-li znám objem a struktura celkové produkce, můžeme určit strukturu a objem konečné produkce i přidané hodnoty, stačí jen dosadit do rovnice (6c). Je-li dána struktura a objem konečné produkce, dostaneme objem a strukturu celkové produkce a přidané hodnoty ze vztahu:

(7a) $$(\mathbf{E} - \mathbf{A})^{-1}\,\mathbf{y} = \mathbf{x}\,,$$

(7b) $$\mathbf{A}_z(\mathbf{E} - \mathbf{A})^{-1}\,\mathbf{y} = \mathbf{z}$$

a souhrnně:

(7c) $$\begin{bmatrix} \mathbf{E} - \mathbf{A} & \mathbf{O} \\ \\ \mathbf{A}_z & \mathbf{O} \end{bmatrix}^{-1} \mathbf{y} = \begin{bmatrix} \mathbf{x} \\ \\ \mathbf{z} \end{bmatrix}.$$

Prvky inverzní matice $(\mathbf{E} - \mathbf{A})^{-1}$ a matice $\mathbf{A}_z(\mathbf{E} - \mathbf{A})^{-1}$ se nazývají koeficienty plných nákladů.

Rozměry meziodvětvových bilancí a jejich agregace

Rozměry meziodvětvových bilancí dnes sestavovaných se značně liší. S rozsáhlými tabulkami se 400—500 obory v prvním kvadrantu kontrastují tabulky malé, členěné na deset nebo i méně odvětví. Volba počtu odvětví vypadá na prvý pohled jednoduše, zdá se, že záleží jen na spolehlivosti statistiky, na účelu bilancování a možnostech výpočetní techniky. Čistě teoreticky můžeme opravdu uvažovat o modelech obsahujících velmi velký počet proměnných. Ve skutečnosti však jsou rozměry meziodvětvových bilancí omezeny tím, že vlastnosti odvětví, výrobních oborů nebo skupin výrobků v praxi nikdy plně neodpovídají třem základním teoretickým předpokladům meziodvětvové analýzy:

a) každý výrobek se vyrábí v jednom odvětví;
b) v každém odvětví se vyrábí jeden výrobek a používá jediné technologie výroby;
c) výrobky různých odvětví jsou nezastupitelné;

Při hlubším třídění lze lépe vyhovět prvým dvěma požadavkům, avšak snadno se naruší třetí pravidlo. Podrobné údaje o struktuře výroby se proto v praxi shrnují do nepříliš velkých meziodvětvových tabulek a ty se často pro různé účely ještě dále zmenšují. Takový postup se nazývá *agregace* meziodvětvové bilance.

Problémy spojené s agregací se v posledních letech intenzívně studují. Byly např. odvozeny podmínky tzv. přípustné agregace (viz např. práce [1], [9] [27]). Jsou však velmi striktní a prakticky nesplnitelné. Pokus odvodit z nich kritérium relativní výhodnosti různých variant agregace najdeme v práci [29]. Podobnou metodu, využívající údajů o rozptylu původních koeficientů kolem koeficientů agregovaných navrhl W. D. Fisher [6]. Na možnost použití korelační závislosti k vyjádření podobnosti struktury nákladové struktury odvětví upozorňuje Kossov [8], na význam relativní stability podílu agregovaných odvětví na celkové a konečné produkci Yamada [16].

Každá z těchto metod klade důraz na jiné vlastnosti meziodvětvové bilance. Shrneme-li jejich výsledky, zjistíme, že při agregaci meziodvětvových bilancí je třeba přihlížet k těmto skutečnostem:

a) podobnosti technologie výroby (podobnosti koeficientů přímých nákladů);
b) podobnosti užitné hodnoty produkce (podobnosti struktury rozdělení produkce);
c) technologické návaznosti odvětví;
d) stabilitě podílu odvětví na celkové a konečné produkci;
e) absolutnímu objemu produkce.

Žádná z metod dosud navržených nepřihlíží současně ke všem uvedeným vlastnostem (které někdy mohou být i v protikladu.) Syntetičtější ukazatel získáme, použijeme-li při řešení problému elementárních poznatků teorie informace.

AGREGACE MEZIODVĚTVOVÝCH BILANCÍ Z HLEDISKA JEJICH INFORMAČNÍHO OBSAHU

Informační obsah meziodvětvových bilancí

Podívejme se na meziodvětvovou bilanci z poněkud nezvyklého hlediska a položme si otázku, v čem záleží její „informační obsah", čím je jeho velikost určována, co ho zvyšuje a co snižuje.

Na první pohled se nabízí souvislost velikosti „informačního obsahu" s rozměry bilance. Z velké tabulky se o ekonomickém systému dozvíme více než z tabulky malé. Zmenší-li se bilance agregací, její „informační obsah" bude zřejmě nerostoucí, zvětší-li se desagregací, „informační obsah" bude zřejmě neklesající. Velikost bilance však není jediným faktorem, ke kterému musíme přihlížet. Srovnáme-li dvě stejně velké tabulky, zjistíme, že jejich „informační obsah" nemusí být stejně veliký. Není lhostejné, je-li objem produkce všech odvětví přibližně stejný nebo převyšuje-li podíl jednoho odvětví podstatně podíl ostatních. Stejně tak záleží na tom, vyskytují-li se odvětví s podobnou technologií výroby nebo podobnou užitnou hodnotou produkce. Obrazně řečeno, záleží na rozlišitelnosti obrazu o struktuře ekonomického systému, na tom, je-li tato struktura, metaforicky řečeno, vykreslena ostrými konturami nebo je-li mlhavá a nevýrazná.

Zatím jsme používali pojmu „informační obsah" bez přesnější definice a chápali ho tak, že se z určitých údajů dovíme více nebo méně o skutečnosti, která nás zajímá. Nyní se ho pokusíme matematicky vyjádřit. Jeho souvislost s určitostí, rozlišitelností struktury meziodvětvové bilance naznačuje, že míru pro vyjádření „informačního obsahu" musíme hledat v teorii informace, tj. ve způsobu, jakým se vyjadřuje entropie.

Velikost entropie se měří Shannonovou mírou entropie, jež se obvykle definuje takto: Mějme množinu n prvků, z nichž každý se vyskytuje s pravděpodobností $p_1, p_2, \ldots, p_n$. Součet pravděpodobností se rovná jedné:

$$\sum_{i=1}^{n} p_i = 1 . \tag{8}$$

Entropií množiny nazýváme veličinu H:

$$H = \sum_{i=1}^{n} - p_i \log p_i . \tag{9}$$

Takto definovaná míra entropie nabývá maximální hodnoty, jsou-li pravděpodobnosti výskytu všech prvků stejné. V takovém případě $H = \log n$ a prvky množiny jsou nerozlišitelné. Minimální, tj. nulové hodnoty veličina H nabývá, vyskytuje-li se jeden prvek s pravděpodobností $p_i = 1$ a prvky ostatní s pravděpodobností nulovou.

Entropie meziodvětvových bilancí

Abychom mohli aplikovat Shannonovu míru entropie na bilanci meziodvětvových vztahů, vyjádříme strukturální vztahy formou, která tuto aplikaci umožňuje. Bilanci zapíšeme nejdříve ve tvaru:*

$$\begin{array}{ccccccccccc}
x_{11} & + \ldots + & x_{1j} & + \ldots + & x_{1n} & = & x_1 & - & y_1 , \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
x_{i1} & + \ldots + & x_{ij} & + \ldots + & x_{in} & = & x_i & - & y_i , \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
x_{n1} & + \ldots + & x_{nj} & + \ldots + & x_{nn} & = & x_n & - & y_n , \\
\hline
z_{11} & + \ldots + & z_{1j} & + \ldots + & z_{1n} & = & z_1 , & & \\
\vdots & & \vdots & & \vdots & & \vdots & & \\
z_{k1} & + \ldots + & z_{kj} & + \ldots + & z_{kn} & = & z_k , & & \\
\vdots & & \vdots & & \vdots & & \vdots & & \\
z_{m1} & + \ldots + & z_{mj} & + \ldots + & z_{mn} & = & z_m , & & \\
\hline
x_1 & + \ldots + & x_j & + \ldots + & x_n & = & x . & & \\
\hline
\end{array} \tag{10}$$

* Autor je zavázán dr. J. Hájkovi za radu a pomoc při řešení tohoto problému.

A pak zavedeme výrazy:

(11a) $$p_{ij} = \frac{x_{ij}}{x},$$

(11b) $$p'_{kj} = \frac{z_{kj}}{x},$$

(12a) $$v_i = \frac{x_i - y_i}{x},$$

(12b) $$v'_k = \frac{z_k}{x},$$

(13) $$w_j = \frac{x_j}{x}.$$

Pro tyto podíly platí:

(14) $$\sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} + \sum_{k=1}^{m} \sum_{j=1}^{n} p'_{kj} = 1,$$

(15) $$\sum_{i=1}^{n} v_i + \sum_{k=1}^{m} v'_k = 1,$$

(16) $$\sum_{j=1}^{n} w_j = 1.$$

Vyhovují tedy vztahu (8), který je podmínkou pro užití Shannonovy míry entropie. Taková úprava meziodvětvové tabulky je ovšem čistě formální a nedává odpověď na otázku, jak pomocí Shannonovy míry entropie vyjádřit rozlišitelnost obrazu o ekonomické struktuře. Cestu k řešení naznačíme nejdříve ekonomickou úvahou a pak ji budeme formulovat matematicky.

Meziodvětvová bilance zobrazuje vzájemný vztah mezi celkovou a konečnou produkcí. Představme si, že ho chceme hlouběji osvětlit a známe jen globální údaje o obou faktorech. Nevíme, jak se na celkové a konečné produkci podílejí jednotlivá odvětví a neznáme ani strukturu nákladů v odvětvích ani strukturu rozdělení jejich produkce. V takovém případě bude analýza závislosti obou ukazatelů velmi neurčitá. Kdybychom z údajů, které máme k dispozici, chtěli sestavit meziodvětvovou bilanci, museli bychom předpokládat, že objem produkce všech odvětví je stejný, a že odvětví mají rovnoměrnou strukturu nákladů i rovnoměrné rozdělení produkce. Taková tabulka by se zřejmě vyznačovala velkou (a jak v dalším uvidíme, maximálně možnou) entropií a obsahovala by „nulovou informaci". Nulovou v tom smyslu, že nic neříká o struktuře daného sytému.

Předpokládejme, že v dalším kroku zjistíme podíl odvětví na celkové a konečné produkci. Nejistota o povaze vzájemného vztahu mezi oběma ukazateli se tím poněkud sníží. Kdybychom podle nových znalostí upravili neurčitou meziodvětvovou bilanci z předchozího případu, korigovali bychom její okrajové podmínky. Struktura vnitřku by však zůstala neznámá a museli bychom o ní nadále předpokládat, že je rovnoměrná, ovšem rovnoměrná v souladu s okrajovými hodnotami. Taková tabulka již obsahuje určitou informaci, skrytou v okrajových podmínkách. Nazveme ji tabulkou s „marginální informací".

Zjistíme-li strukturu nákladů v odvětvích a strukturu rozdělení jejich produkce, neurčitý obraz o struktuře ekonomického systému se podstatně zpřesní a tabulka se změní v normální meziodvětvovou bilanci.

Přechod do naprosté nerozlišitelnosti tabulky s „nulovou informací" k větší rozlišitelnosti obrazu o struktuře ekonomického systému v meziodvětvové bilanci lze vyjádřit pomocí Shannonovy míry entropie. Jednotlivé typy tabulek můžeme charakterizovat takto:

Tabulkou s nulovou informací rozumíme tabulku, v níž platí rovnice:

(17a) $$p_{ij} = p'_{kj} = \frac{1}{n(n + m)},$$

(17b) $$v_i = v'_k = \frac{1}{n + m},$$

(17c) $$w_j = \frac{1}{n}.$$

Míra entropie nabývá maximální možné hodnoty a je:

(18) $$H^0 = \log n(n + m).$$

Veličinu H^0 můžeme také považovat za číselnou charakteristiku množiny meziodvětvových bilancí, pro které jsou známa čísla m a n (počet odvětví a počet složek přidané hodnoty).

V tabulce s marginální informací známe hodnoty v_i, v'_k a w_j. Ostatní hodnoty jsou dány vztahem:

(19a) $$p_{ij} = v_i w_j,$$

(19b) $$p'_{kj} = v'_k w_j.$$

Její entropie je:

(20) $$\begin{aligned} H^m &= -\sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} \log p_{ij} - \sum_{k=1}^{m} \sum_{j=1}^{n} p'_{kj} \log p'_{kj} = \\ &= -\sum_{i=1}^{n} \sum_{j=1}^{n} v_i w_j \log v_i w_j - \sum_{k=1}^{m} \sum_{j=1}^{n} v'_k w_j \log v'_k w_j = \\ &= -\sum_{j=1}^{n} v_i \log v_i - \sum_{k=1}^{m} v'_k \log v'_k - \sum_{j=1}^{n} w_j \log w_j. \end{aligned}$$

 Ukazatel H^m lze též interpretovat jako charakteristiku množiny meziodvětvových bilancí s danou strukturou celkové produkce a danou strukturou nákladů na společenský produkt.

Třetí typ tabulky nazveme tabulkou s normální informací. Je totožná s bilancí meziodvětvových vztahů, popsanou soustavou rovnic (10). Od tabulky s marginální informací se liší tím, že existuje alespoň jeden z indexů i ($i = 1, 2, \ldots, n$) a j ($j = 1, 2, \ldots, n$) takový, že platí:

(21a) $$p_{ij} \neq v_i w_j \, ,$$

nebo alespoň jeden z indexů k ($k = 1, 2, \ldots, m$) a j ($j = 1, 2, \ldots, n$), že platí:

(21b) $$p'_{kj} \neq v_k w_j \, .$$

Její entropie je dána výrazem:

(22) $$H = -\sum_{i=1}^{n} \sum_{j=1}^{a} p_{ij} \log p_{ij} - \sum_{k=1}^{n} \sum_{j=1}^{n} p'_{kj} \log p'_{kj} \, .$$

Míra informačního obsahu meziodvětvové bilance

Entropie tabulek s nulovou, marginální a normální informací ještě nevyjadřuje, kolik informace je v meziodvětvové bilanci obsaženo. Informační obsah meziodvětvové bilance vyplývá z hlavního cíle meziodvětvového bilancování. Při něm nás nezajímají rozdíly podílů odvětví na celkové a konečné produkci (ty lze zjistit daleko jednodušeji), ale především odlišnosti jejich nákladové struktury a struktury rozdělení produkce. Čím výraznější obraz o vnitřní struktuře ekonomického systému bilance poskytují, tím je jejich „informační obsah" větší. Proto nás především zajímá rozdíl entropie tabulek s informací marginální a normální. Míru množství informace, která vyjadřuje velikost informačního obsahu meziodvětvové bilance, pak definujeme jako rozdíl entropie tabulek s marginální a normální informací:

(23) $$I = H^m - H \, .$$

Takto definovaná míra množství informace má tyto vlastnosti:

a) rovná se nule, používají-li všechna odvětví stejné technologie výroby a rozděluje-li se produkce všech odvětví ve stejných proporcích. V tomto případě platí vztahy (19a) a (19b) a lze ukázat, že struktura nákladů všech odvětví i struktura rozdělení jejich produkce je ve všech odvětvích stejná;

b) monotónně vzrůstá se zvětšováním rozměrů bilance při její desagregaci a monotónně klesá při jejím zmenšování agregací.

Snížení míry množství informace lze vypočítat pro libovolnou agregaci dvou, tří nebo obecně libovolného počtu n-tic odvětví. Analýza složitějších případů je však velmi obtížná a prakticky, při posuzování různých variant agregace, nerealizovatelná ani s pomocí samočinného počítače. Proto v dalším ukážeme jen nejelementárnější případ agregace dvojice odvětví.

Sloučíme-li v bilanci o n odvětvích dvě odvětví, označená symboly g a h, do odvětví g^* (tj. snížíme-li počet odvětví na $n - 1$), pak platí, že míra množství informace se zmenší o veličinu D, která je dána výrazem:

(24)
$$D = \sum_{i=g,h} \sum_{j=g,h} p_{ij} \log \frac{p_{ij}}{\sum\limits_{i=g,h} \sum\limits_{j=g,h} p_{ij}} +$$
$$+ \sum_{\substack{i=1 \\ (i \neq g,h)}}^{n} p_{ig} \log \frac{p_{ig}}{p_{ig} + p_{ih}} + \sum_{\substack{i=1 \\ (i \neq g,h)}}^{n} p_{ih} \log \frac{p_{ih}}{p_{ig} + p_{ih}} +$$
$$+ \sum_{k=1}^{m} p'_{kg} \log \frac{p'_{kg}}{p'_{kg} + p'_{kh}} \sum_{k=1}^{m} p_{kh} \log \frac{p'_{kh}}{p'_{kg} + p'_{kh}} +$$
$$+ \sum_{\substack{j=1 \\ (j \neq g,h)}}^{n} p_{gj} \log \frac{p_{gj}}{p_{gj} + p_{hj}} + \sum_{\substack{j=1 \\ (j \neq g,h)}}^{n} p_{hj} \log \frac{p_{hj}}{p_{gj} + p_{hj}} -$$
$$- v_g \log \frac{v_g}{v_g + v_h} - v_h \log \frac{v_h}{v_g + v_h} -$$
$$- w_g \log \frac{w_g}{w_g + w_h} - w_h \log \frac{w_h}{w_g + w_h} .$$

Lze dokázat (důkazy pro nedostatek místa neuvádíme, jde však o vztahy, které jsou snadno pochopitelné), že snížení míry množství informace závisí na třech faktorech:

a) podobnosti nákladové struktury slučovaných odvětví, tedy na podobnosti používané technologie výroby;

b) podobnosti struktury rozdělení produkce slučovaných odvětví, tedy na podobnosti užití produkce slučovaných odvětví;

c) objemu produkce slučovaných odvětví.

Některé výsledky zkušebních propočtů naznačují, že na výši D působí i návaznost odvětví v technologickém procesu. Platnost tohoto vlivu se však zatím nepodařilo obecně prokázat.

Nejmenší snížení míry množství informace při agregaci bilance tedy nastane, spojíme-li odvětví nevelká, s podobnou technologií výroby a podobným užitím produkce. Tyto vlastnosti míry množství informace plně odpovídají požadavkům na agregaci odvětví v meziodvětvové bilanci.

Postupné slučování dvojic odvětví nemusí obecně vést k optimálnímu výsledku, určité seskupování dvojic nezaručuje optimální výběr pětic, šestic či n-tic odvětví. Přesto se domníváme, že při takovém postupu nebudou praktické výsledky daleko od optima. Proto byl pro postupné slučování dvojic odvětví na základě míry množství informace sestaven program* pro samočinný počítač National Elliot a provedeny prvé zkušební propočty pro nevelké meziodvětvové bilance.

Ve výpočtu se opakují tři základní operace:

a) Výpočet hodnot p_{ij}, p'_{kj}, v_i, v'_k a w_j podle vzorců (11)–(13);

b) výpočet poklesu míry množství informace podle vzorce (24) pro všechny varianty agregace dvojic odvětví v prvém, resp. agregace složek přidané hodnoty ve třetím kvadrantu bilance;

c) vyhledání minimální hodnoty ukazatele D, tedy určení dvojice odvětví, jejichž agregace způsobí nejmenší pokles míry množství informace.

V dalším kroku opravíme hodnoty p_{ij}, p'_{kj}, v_i, v'_k a w_j pro nové sloučené odvětví. Stejně tak opravíme hodnoty D. Následuje výběr minimálního D a celý postup se opakuje.

(Došlo dne 27. června 1964.)

LITERATURA

[1] Ara K.: The Aggregation Problem in Input-Output Analysis. Econometrica (1959), č. 2, 257—262.

[2] Balderston J. B., Whitin T. M.: Aggregation in the Input-Output Model, sborník Economic Activity Analysis, New York 1949, 79—102.

[3] Barna T.: Classification and Aggregation in Input-Output Analysis, sborník The Structural Interdependence of the Economy, New York 1957, 173—186.

[4] Czechowski T.: Kilka uwag o agregacji w statycznym modelu W. Leontiewa. Przeglad statystyczny (1960), č. 4.

[5] Evans W. D.: Input-Output Computations. Sborník The Structural Interdependence of the Economy, New York 1957, 51—102.

[6] Fisher W. D.: Criteria for Aggregation in Input-Output Analysis. The Review of Economics and Statistics (1958), č. 3.

[7] Яглом А. М., Яглом И. М.: Вероятность и теория информации, Москва 1960.

[8] Коссов В. В.: Возможные решения проблемы агрегирования в межотраслевом балансе, Вопросы экономики 1963, № 6.

* Autorem programu je J. Sehnal, prom. mat.

[9] Malinvaud E.: Aggregation Problems in Input-Output Models. Sborník The Structural Interpendence of the Economy, 1957, 187—202.
[10] Morgenstern O.: Experiment and Large Scale Computation in Economics. Sborník Economic Activity Analysis, New York 1949, 483—549.
[11] Morishima M., Seton F.: Aggregation in Leontief Matrices and Labour Theory of Value. Econometrica (1941), č. 4.
[12] Popelka M., Sekerka B.: K problému agregace v bilanci cenových vlivů. Politická ekonomie (1963), č. 7, 544—556.
[13] Skolka J.: Agregace v meziodvětvových bilancích. Sborník statistika a demografie, Praha 1963, 67—86.
[14] Stone R.: Agregacja v modelech Input-Output. Przeglad statystyczny, (1962), č. 1, 25—28.
[15] Theil H.: Linear Aggregation in Input-Output Analysis, Econometrica (1957), č. 1, 111—122.
[16] Yamada I.: Theory and Application of Interindustry Analysis. Tokyo 1961.

SUMMARY

The Use of Information Measure for Aggregation of Input-Output Tables

JIŘÍ SKOLKA

For finding an appropriate aggregation of input-output tables which describe the structure of economic systems several techniques have been proposed. In the present paper the use of Shannon's entropy rate is demonstrated. The starting point is the concept of the information content of an input-output table. It depends both on the number of industries and on the discernability of the economical structure. We proceed in such a way that the input-output table is modified to the form (10) and the expressions (11)–(13) are introduced. Then the entropy rate for the table (10) and for the theoretical table given by conditions (19) is determined. Their difference is the measure of information. If we aggregate the input-output table in such a way that the corresponding measure of information decreases as little as possible it can be proved that we aggregate industries with similar production functions, with similar structure of distribution and with small production volume. This fully conforms to the economical requirements of aggregation.

Inž. Jiří Skolka, CSc., Ekonomický ústav ČSAV, Ekonomicko-matematická laboratoř, Politických vězňů 7, Praha 1.