Andrej Pázman

Zachovanie informácie pri redukcii štatistických modelov merania

Matematicko-fyzikálny časopis, Vol. 15 (1965), No. 1, 81--89

Persistent URL: `http://dml.cz/dmlcz/126389`

ZACHOVANIE INFORMÁCIE PRI REDUKCII ŠTATISTICKÝCH MODELOV MERANIA

ANDREJ PÁZMAN, Bratislava

V práci sa analyzuje možnosť redukcie údajov výstupu meracieho systému pri zachovaní informácie. Odvodené všeobecné vzťahy sú použité najprv na analýzu lineárneho modelu merania s gaussovskými chybami merania. Tento model zahrňuje v sebe priame i nepriame, závislé i nezávislé merania vyrovnávacieho počtu (teórie chýb). Ďalej sa analyzuje model opakovaných meraní s výrazne nesymetrickými aditívnymi chybami merania.

Základným predpokladom každého merania je určitá znalosť fyzikálnych procesov, podľa ktorých sa správa meraný objekt i merací systém. Ak však zväčšujeme požiadavky presnosti merania alebo budujeme značne zložité meracie systémy, skôr či neskôr si musíme všímať vplyvy, priebeh ktorých nevieme vystihnúť deterministicky. Obyčajne tieto vplyvy vieme vystihnúť aspoň štatisticky, prípadne obmedziť interval ich pôsobenia a pod.

Zameriame sa na také meracie systémy, ktoré merajú „súčasne" k meraných veličín $x_1, \ldots, x_k$ a ako predbežný výsledok merania poskytujú m hodnôt $(m \geqq k)$ $y_1, \ldots, y_m$, ktoré budeme stručne nazývať výstupom meracieho systému. Predpokladáme, že z analýzy meracieho systému je nám známa pravdepodobnostná súvislosť medzi výstupom a vstupom meracieho systému daná hustotou pravdepodobnosti

$$f_1(Y/X),$$

kde Y, X sú vektory so zložkami $y_1, \ldots, y_m$ a $x_1, \ldots, x_k$.

Z výstupu meracieho systému Y nemôžeme určiť presné hodnoty meraných veličín. V princípe môžeme však určiť k-rozmerné priestory spoľahlivosti, ktoré s vopred predpísanou pravdepodobnosťou P pokrývajú neznáme hodnoty meraných veličín.

Konštrukcia priestorov spoľahlivosti je jednoduchšia, ak $m = k$. Je ďalej obťažné priamo porovnať medzi sebou rôzne meracie systémy, ktoré slúžia na meranie tých istých veličín a ktoré majú rôzny počet (m) výstupných hodnôt $y_1, \ldots, y_m$. Je preto účelné, ako prvý krok pri spracovaní predbežných výsledkov merania a pri porovnávaní rôznych meracích systémov vykonať redukciu výstupu nameraných hodnôt. m hodnôt $y_1, \dots, y_m$, ktoré sú výberom z náhodného súboru s hustotou pravdepodobnosti $f_1(Y/X)$, potrebujeme nahradiť menším počtom k hodnôt $z_1, \dots, z_k$, ktoré sú výberom z náhodného súboru s hustotou pravdepodobnosti $f_2(Z/X)$. (Obmedzíme sa na také prípady, keď hustota $f_2(Z/X)$ existuje.)

Vo výstupe meracieho systému Y je obsiahnuté určité množstvo informácie o meraných veličinách X. Vyžadujeme, aby sa toto množstvo informácie zachovávalo pri redukcii vyjadrenej transformáciou $Z = Z(Y)$.

V prípade, že hodnota vektora X je výberom z náhodného súboru s rozdelením pravdepodobnosti p_X, množstvo informácie (v Shannonovom zmysle) o veličine X obsiahnuté vo veličine Y je dané vzťahom (pozri [2])

$$I(X, Y) = M \left\{ \log \frac{f_1(Y/X)}{\int_{E_k} f_1(Y/X) \, \mathrm{d}p_X} \right\}. \tag{1}$$

kde $M\{\}$ je symbol strednej hodnoty a E_k označuje k-rozmerný euklidovský priestor. Podobne informácia o veličine X obsiahnutá vo veličine Z je daná výrazom:

$$I(X, Z) = M \left\{ \log \frac{f_2(Z/X)}{\int_{E_k} f_2(Z/X) \, \mathrm{d}p_X} \right\}. \tag{2}$$

Podmienka zachovania informácie pri transformácii $Z = Z(Y)$ je

$$\Delta I = I(X, Y) - I(X, Z) = 0. \tag{3}$$

I keď pripustíme, že výskyt rôznych hodnôt meraných veličín je vzhľadom na experimentátora náhodný, nepoznáme rozdelenie pravdepodobnosti p_X charakterizujúce túto náhodnosť. Preto budeme vyžadovať, aby vzťah (3) bol splnený pre všetky možné (fiktívne) rozdelenia pravdepodobnosti p_X.

Lema 1. *Pre každé číslo u: $0 < u < \infty$ platí:*

$$\log u \geqq 1 - 1/u, \tag{4}$$

pričom rovnosť vo vzťahu (4) platí vtedy a len vtedy, keď $u = 1$.

Dôkaz. Vezmime funkciu

$$g(u) = \log u + 1/u - 1.$$

Pomocou prvej a druhej derivácie zistíme, že funkcia $g(u)$ má jediné minimum pre $0 < u < \infty$, a to v bode $u = 1$, pričom $g(1) = 0$.

Veta 1. *Úbytok informácie ΔI je vždy nezáporný a transformácia*

$$Z = Z(Y)$$

zachováva informáciu o veličine X ($\Delta I = 0$ pre každé p_X) vtedy a len vtedy, keď existuje nejaká funkcia $\Phi(Y)$ závislá iba od Y taká, že

$$\frac{f_1(Y/X)}{f_2[Z(Y)/X]} = \Phi(Y) \tag{5}$$

pre všetky $X \in E_k$ a pre skoro všetky Y.

Dôkaz. Zvoľme si ľubovoľné p_X. Označme

$$\varphi_1(Y, p_X) = \int_{E_k} f_1(Y/X)\mathrm{d}p_X; \tag{6}$$

$$\varphi_2(Y, p_X) = \int_{E_k} f_2[Z(Y)/X]\mathrm{d}p_X. \tag{7}$$

Dosadením (1) a (2) do vzťahu pre úbytok informácie dostaneme:

$$\Delta I = I(X, Y) - I(X, Z) = M\left\{\log\frac{f_1(Y/X) \cdot \varphi_2(Y, p_X)}{f_2[Z(Y)/X] \cdot \varphi_1(Y, p_X)}\right\}. \tag{8}$$

Použitím lemy 1 a vzťahu (8) dostaneme:

$$\Delta I \geq M\left\{1 - \frac{f_2 \cdot \varphi_1}{f_1 \cdot \varphi_2}\right\} = 1 - \int_{E_n}\int_{E_k} \frac{f_2 \cdot \varphi_1}{f_1 \cdot \varphi_2}\,\mathrm{d}p_X f_1 \mathrm{d}Y = \tag{9}$$

$$= 1 - \int_{E_n}\left[\frac{\varphi_1}{\varphi_2}\int_{E_k} f_2\,\mathrm{d}p_X\right]\mathrm{d}Y.$$

Rozpísaním φ_1 a φ_2 vo vzťahu (9) dostaneme

$$\Delta I \geq 1 - \int_{E_k}\left[\int_{E_n} f_1(Y/X)\mathrm{d}Y\right]\mathrm{d}p_X = 0. \tag{10}$$

Podľa lemy 1 rovnosť vo vzťahu (10) platí vtedy a len vtedy, keď pre skoro všetky Y (v zmysle Lebesgueovej miery μ) a skoro všetky X (v zmysle miery p_X) platí:

$$\frac{f_1(Y/X)}{f_2[Z(Y)/X]} = \frac{\varphi_1(Y, p_X)}{\varphi_2(Y, p_X)}. \tag{11}$$

Nech platí (5). Dosadením do (11) podľa vzťahov (5), (6) a (7) potvrdíme rovnosť vo vzťahu (11) pre každé p_X. Dostatočnosť tvrdenia vety je dokázaná.

Nech naopak platí (11) pre všetky možné p_X. Potom zrejme vzťah (11) musí platiť pre všetky $X \in E_k$. Ľavá strana rovnosti (11) nie je závislá od p_X, preto ani pomer φ_1/φ_2 nie je závislý od p_X. Veta je dokázaná.

Poznámka. Podmienka (5) je rovnaká ako podmienka dostatočnosti štatistík (v zmysle teórie R. Fishera). V práci však nejde o určovanie štatistík (bodových odhadov), ale o redukciu a o určenie priestorov spoľahlivosti.

Ku každej hodnote Y výstupu meracieho systému priradíme merateľnú množinu $B(Y)$, prvkami ktorej sú vektory X. Ak je pre každé pevné X pravdepodobnosť namerania výstupu Y takého, že $X \in B(Y)$ je aspoň P, potom množiny $B(Y)$ tvoria P-priestory spoľahlivosti (pozri [1]). Pretože konštrukcia priestorov spoľahlivosti nie je jednoznačná, vyžadujeme, aby „objem" (Lebesgueova miera μ) množín $B(Y)$ bol minimálny. Minimálne P-priestory spoľahlivosti $B'(Z)$ môžeme konštruovať aj pre redukovaný výstup Z.

Veta 2. *Nech transformácia $Z = Z(Y)$ zachováva informáciu. Nech merateľné množiny $B'(Z) \subset E_k$ majú nasledujúce vlastnosti:*

a) *Množiny $B'(Z)$ (závislé od náhodného vektora Z) sú P-priestormi spoľahlivosti pre určenie skutočnej hodnoty vektora X.*

b) *Ak $X_1 \in B'(Z)$, $X_2 \notin B'(Z)$, tak $f_2(Z/X_1) > f_2(Z/X_2)$.*

c) *Objem (miera μ) priestorov spoľahlivosti $B'(Z)$ nezávisí od Z.*

Potom platí:

1. *Množiny $B'(Z)$ sú jednoznačne určenými (až na množinu nulovej miery μ) minimálnymi P-priestormi spoľahlivosti priradenými k redukovanému výstupu Z.*

2. *Množiny*

$$B(Y) = B'[Z(Y)] \tag{12}$$

sú jednoznačne určenými (až na množinu miery nula) minimálnymi P-priestormi spoľahlivosti priradenými k výstupu Y.

Dôkaz. Označme:

$$A(X) = \{Z : X \in B'(Z)\} . \tag{13}$$

Pre ľubovoľnú konštrukciu priestorov spoľahlivosti $B'(Z)$ musí platiť:

$$p\{A(X)/X\} = \int_{A(X)} f_2(Z/X) \mathrm{d}Z = P . \tag{14}$$

Zo vzťahu (14) vyplýva:

$$\lim_{a \to \infty} \frac{1}{\mu[G(a)]} \int_{G(a)} p\{A(X)/X\} \,\mathrm{d}X = P , \tag{15}$$

kde $G(a)$ je k-rozmerná guľa o polomere a. Na základe vzťahu (13) a rozpísaním vzťahu (15) dostaneme:

$$(16) \qquad P = \lim_{a\to\infty} \frac{1}{\mu[G(a)]} \int\limits_{G(a)} \left[\int\limits_{A(X)} f_2(Z/X)\, \mathrm{d}Z \right] \mathrm{d}X =$$

$$= \lim_{b\to\infty} \frac{1}{\mu[G(b)]} \int\limits_{G(b)} \left[\int\limits_{B'(Z)} f_2(Z/X)\, \mathrm{d}X \right] \mathrm{d}Z\,.$$

Posledný vzťah musí byť splnený pre ľubovoľnú konštrukciu P-priestorov spoľahlivosti. Z neho a z podmienky b) vyplýva, že každá konštrukcia P-priestorov spoľahlivosti odlišná od priestorov spoľahlivosti spĺňajúcich podmienky a), b), c), na množine nenulovej miery, má nutne väčší „objem" (mieru).

2. Z podmienky (5) a zo vzťahu (12) vyplýva, že hustota pravdepodobnosti $f_1(Y/X)$ a množiny $B(Y)$ spĺňajú predpoklady a), b), c), a preto $B(Y)$ dané vzťahom (12) sú jednoznačne určenými minimálnymi P-priestormi spoľahlivosti priradenými výstupu Y.

Poznámka. Z vety 2 vyplýva: Ak $Z = Z(Y)$ a $W = W(Y)$ sú dve transformácie zachovávajúce informáciu, obe vedú k tým istým priestorom spoľahlivosti.

Gaussovský lineárny model merania

Uvažujme také merania, pri ktorých medzi meranými veličinami a výstupom meracieho systému je lineárna závislosť známa z konštrukcie meracieho systému

$$(17) \qquad y_i^* = \sum_{j=1}^{k} s_{ij} x_j + a_i \qquad (i = 1, \ldots, m)\,,$$

pričom merania sú sprevádzané náhodnými aditívnymi chybami merania s gaussovským rozložením pravdepodobnosti, t. j.

$$(18) \qquad y_i = y_i^* + n_i \qquad (i = 1, \ldots, m)\,,$$

n_i je hodnota i-tej chyby merania. Ak spojíme vzťahy (17) a (18) a zapíšeme v maticovom tvare, dostaneme:

$$(19) \qquad Y = SX + A + N$$

a hustota podmienenej pravdepodobnosti má tvar[1]

$$(20) \quad f_1(Y/X) = \frac{1}{(2\pi)^{\frac{1}{2}m}\,(\det K)^{\frac{1}{2}}} \exp\{-\tfrac{1}{2}[Y - SX - A]^T K^{-1} [Y - SX - A]\}\,,$$

[1] Exponentom T označujeme transpozíciu, exponentom -1 označujeme inverziu matice.

kde K je matica kovariancií chýb merania. V prípade, že je diagonálna, uvažovaný model je model nepriamych nezávislých meraní známy z vyrovnávacieho počtu. Pretože priame merania sú špeciálnym prípadom nepriamych a závislé merania sa pomocou elementov dajú previesť na nezávislé, týka sa nasledujúca analýza aj priamych i nepriamych, závislých i nezávislých meraní vyrovnávacieho počtu.

Lema 2. *Ak*

$$Z = LY + B\,,$$

kde vektorová náhodná veličina Y má hustotu pravdepodobnosti (20) a L je matica typu k/n hodnosti k, potom existuje hustota pravdepodobnosti $f_2(Z/X)$ a má tvar:

$$f_2(Z/X) = \frac{1}{(2\pi)^{\frac{1}{2}m}(\det R)^{\frac{1}{2}}} \exp\left\{-\tfrac{1}{2}[Z - M]^T R^{-1}[Z - M]\right\}, \tag{21}$$

kde stred rozdelenia M a kovariančná matica R sú dané vzťahmi:

$$M = M(X) = LSX + LA + B; \tag{22}$$

$$R = LKL^T. \tag{23}$$

Dôkaz pomocou charakteristickej funkcie pozri v [1].

Veta 3. V gaussovskom lineárnom modeli merania redukcia pomocou lineárnej transformácie

$$Z = LY + B \tag{24}$$

zachováva informáciu vtedy a len vtedy, keď matica L má tvar

$$L = D(S^T K^{-1} S)^{-1} S^T K^{-1}\,, \tag{25}$$

kde D je ľubovoľná nesingulárna k-rozmerná štvorcová matica.

Dôkaz. Do vzťahu (5) dosadíme podľa (20) a (21) a zlogaritmujeme. Dostaneme:

$$(Y^T - A^T)K^{-1}SX - \tfrac{1}{2}X^T S^T K^{-1} SX - (Z^T(Y) - A^T L^T - B^T)R^{-1}LSX + \tfrac{1}{2}X^T S^T L^T R^{-1} LSX - \log \Phi(Y) = 0\,. \tag{26}$$

Podľa vety 1 transformácia (24) zachováva informáciu vtedy a len vtedy, keď existuje taká funkcia $\Phi(Y)$, že vzťah (26) je splnený pre všetky X a pre skoro všetky Y. Dosadením za $\Phi(Y)$:

$$\Phi(Y) = \frac{f_1(Y/X_1)}{f_2[Z(Y)/X_1]}$$

pri pevne zvolenom vektore X_1 a dosadením za R^{-1}, $Z(Y)$ a L podľa (23), (24) a (25) do ľavej strany vzťahu (26) preveríme platnosť rovnosti (26).

Nech naopak platí rovnosť (26). Kvadratická forma (26) (premennými sú zložky vektora X) je rovná nule len vtedy, keď

$$(Y^T - A^T)K^{-1}S = (Y^T - A^T)L^T R^{-1} LS \tag{27}$$

pre skoro všetky Y. Vzťah (27) môže byť splnený len vtedy, keď

$$K^{-1}S = L^T R^{-1} LS\,. \tag{28}$$

Matica $K^{-1}S$ má hodnosť k. Preto aj matica LS je nesingulárna a existuje matica $(LS)^{-1}$.

Ak vynásobíme vzťah (28) zľava maticou $L^T(S^T L^T)^{-1}S^T$, dostaneme:

$$L^T(S^T L^T)^{-1}S^T K^{-1}S = L^T R^{-1} LS\,. \tag{29}$$

Porovnaním so vzťahom (28) dostaneme:

$$K^{-1}S = L^T(S^T L^T)^{-1}S^T K^{-1}S\,. \tag{30}$$

Transpozíciou a úpravou dostaneme:

$$L = LS(S^T K^{-1}S)^{-1}S^T K^{-1}\,. \tag{31}$$

Matica LS je nesingulárna a je preto jednou z matíc D vo vzťahu (25). Veta je dokázaná.

Existuje celá trieda lineárnych transformácií zachovávajúcich informáciu, ktoré, pravda, podľa vety 2 vedú k tej istej konštrukcii minimálnych priestorov spoľahlivosti. V ďalšom preto vystačíme s transformáciou:

$$Z_v = (S^T K^{-1}S)^{-1}S^T K^{-1}(Y - A)\,. \tag{32}$$

Ak je matica K^{-1} diagonálna, zodpovedá vzťah (32) riešeniu normálnych rovníc vyrovnávacieho počtu, pričom matica K^{-1} zodpovedá matici váh. Veta dopĺňa preto interpretáciu riešenia normálnych rovníc: riešenie normálnych rovníc zachováva všetku informáciu o meraných veličinách $x_1, \ldots, x_k$.

Pre každú predpísanú pravdepodobnosť P určíme číslo r_p vzťahom

$$p\{t : t \leq r_P^2\} = P\,, \tag{33}$$

kde t je náhodná veličina s χ^2 rozdelením pravdepodobnosti. Potom elipsoid:

$$B'(Z) = \{X : [X^T - Z_v^T]S^T K^{-1}S[X - Z_v] \leq r_P^2\} \tag{34}$$

spĺňuje podmienky a), b), c), vety 2, a je preto minimálnym P-priestorom spoľahlivosti priradeným redukovanému výstupu Z_v. Dosadením (32) do (34) dostaneme hľadané elipsoidy spoľahlivosti priradené výstupu Y.

Rôzne meracie systémy (o nerovnakom počte výstupných hodnôt) môžeme porovnať pomocou kovariančnej matice R_v redukovaného výstupu Z_v. Tvar tejto matice dostaneme aplikovaním lemy 2 na vzťah (32). Dostaneme:

$$R_v = (S^T K^{-1}S)^{-1}\,. \tag{35}$$

Môžeme porovnávať aj pomocou objemu elipsoidu spoľahlivosti (34):

$$B'(Z_v) = \mu\,[G(r_p)] \left[\frac{1}{\det\,(S^T K^{-1} S)}\right]^{\frac{1}{2}} \tag{36}$$

($G(r_p)$ je k-rozmerná guľa).

Výraz:

$$\det S^T K^{-1} S = \det R_v^{-1} \tag{37}$$

je vhodnou mierou presnosti merania.

Nesymetrický model merania

Meriame jednu veličinu x (skalárnu). Meranie je sprevádzané aditívnymi nesymetrickými chybami merania n_i s hustotou pravdepodobnosti

$$f_o(n_i) = \begin{cases} a \exp\{-an_i\} & \text{pre } n_i \geqq 0\,, \\ 0 & \text{pre } n_i < 0\,, \end{cases} \tag{38}$$

kde a je kladná konštanta.

Merania m-krát opakujeme (nezávisle) a dostaneme výstupné hodnoty $y_1, \ldots, y_m$ závislé od meranej veličiny podľa vzťahov:

$$y_i = x + n_i \qquad (i = 1, \ldots, m)\,. \tag{39}$$

Zo vzťahov (38) a (39) dostaneme:

$$f_1(Y/x) = \begin{cases} a^m \exp\{-a(\sum_i y_i - mx)\} & \text{pre } y_i \geqq x \ (i = 1, \ldots, m); \\ 0, & \text{ak } y_i < x \text{ aspoň pre jedno } i. \end{cases} \tag{40}$$

Redukciu vykonáme pomocou transformácie:

$$z = \min_i y_i\,. \tag{41}$$

Zo vzťahu (41) vyplýva:

$$p\,\{z : z > z_0\} = \prod_i p\{y_i : y_i > z_0\} \tag{42}$$

a zo vzťahov (42) a (38) dostaneme:

$$f_2(z/x) = \begin{cases} ma \exp\{-ma\,(z - x)\} & \text{pre } z \geqq x, \\ 0 & \text{pre } z < x. \end{cases} \tag{43}$$

Porovnaním hustôt pravdepodobností (40) a (43) zistíme, že spĺňajú podmienku (5) vety 1, t. j. transformácia (41) zachováva informáciu. Pre ľubovoľné P určíme minimálne priestory spoľahlivosti priradené k redukovanému výstupu z zo vzťahu:

$$P = \int_{x}^{x-(1/ma)\log(1-P)} ma \exp\{-ma(v-x)\}\, \mathrm{d}v \,. \tag{44}$$

Označíme ako predtým množiny bodov z spĺňajúce vzťah (44)

$$A(x) = \{z : x < z < x - \frac{1}{ma} \log(1-P)\} \tag{45}$$

a hľadané priestory spoľahlivosti podľa vety 2 sú intervaly:

$$B(Y) = \{x : (\min_i y_i) + \frac{1}{ma} \log(1-P) < x < \min_i y_i\} \,. \tag{46}$$

Vzťah (46) umožňuje priamo z neredukovaného výstupu určiť potrebné intervaly spoľahlivosti.

LITERATÚRA

[1] Cramér H., *Mathematical Methods of Statistics*, Princeton 1946.

[2] Добрушин Р. Л., *Общая формулировка основной теоремы Шеннона в теории информации*, Успехи математических наук *14* (1959), 6, 3—104.

Došlo 20. 7. 1963.

ČSAV, Ústav teórie merania Slovenskej akadémie vied, Bratislava

СОХРАНЕНИЕ ИНФОРМАЦИИ ПРИ СОКРАЩЕНИИ СТАТИСТИЧЕСКИХ МОДЕЛЕЙ ИЗМЕРЕНИЯ

Андрей Пазман

Резюме

В работе анализируется возможность группировки показаний выхода измерительной системы при одновременном сохранении информации. Полученные общие отношения используются сначала для анализа линейной модели измерения с гауссовскими ошибками измерения. В эту модель включены прямые и косвенные, зависимые и независимые измерения, известные из теории ошибок. Общие соотношения используются в дальнейшем для анализа одной модели повторяемых измерений с значительной асимметрией аддитивных ошибок измерения.