# ORTOGONALITATEA – CARACTERISTICĂ A CALITĂȚII BAZEI DE MODELE ECONOMICE

**Ion Ivan**
**Adrian Vişoiu**
**Marius Popa**
*marius.popa@ie.ase.ro*

*Academia de Studii Economice, Bucureşti*

**Rezumat:** Lucrarea defineşte conceptul de model economic şi prezintă principalele caracteristici pe care acesta le îndeplineşte. Sunt trecute în revistă categorii de modele şi procesul de constituire a bazelor de modele. Se evidențiază procesul de analiză a ortogonalității modelelor şi constituirea de clase de modele în funcție de gradul de ortogonalitate a acestora. Testarea şi aplicarea modelelor se realizează pe baza unor seturi de date, pentru care se efectuează, de asemenea, o analiză a ortogonalității lor.

**Cuvinte cheie:** model economic, bază de modele, ortogonalitate, set de date.

## 1. Modele economice

Există numeroase tipuri de modele economice. Unele dintre ele au o singură ecuație, recunoscute de cele mai multe ori sub denumirea de indicatori, şi sunt altele cu mai multe ecuații, modelele trebuind să aibă o serie de proprietăți.

*Proprietatea de senzitivitate* are în vedere variația rezultatelor obținute prin aplicarea modelului la variația valorilor variabilelor de intrare a modelului.

*Caracterul necatastrofic* se referă la absența unor valori particulare, care fac imposibilă obținerea unei valori a variabilei endogene. De exemplu: anularea numitorului, obținerea unei valori negative pentru argumentul funcțiilor radical şi logaritmică etc.

Un model este *necompensatoriu* atunci când, la variații simultane ale unor variabile exogene, nu lipseşte variația variabilei endogene.

Procesul de dezvoltare a modelelor are ca obiectiv realizarea, într-o măsură cât mai mare, a celor trei cerințe. În literatura de specialitate, s-a demonstrat că este imposibilă realizarea de modele care să satisfacă simultan, la nivel maxim, cele trei proprietăți prezentate anterior.

Construirea de modele economice are ca obiectiv analiza corelațiilor dintre factorii de influență şi variabilele rezultative, din care rezultă forme analitice ale dependențelor identificate.

Există o mare diversitate de modele economice, fiecare din ramurile ştiințelor economice rezervându-le spații de prezentare deosebit de largi.

**Modelele directe** sunt rezultatul percepției nemijlocite a variabilelor asociate factorilor de influențe. Dacă se consideră un proces $P$ căruia i se asociază variabila rezultativă $y$ şi care este determinat în evoluția sa de factorii $F_1, F_2, \ldots, F_N$ cărora li se asociază variabilele $X_1, X_2, \ldots, X_N$, un model direct are forma: $y = \sum_{i=1}^{N} a_i X_i$ unde $a_i$ este un coeficient de contribuție, definit pe mulțimea $\{-1,1\}$.

Astfel, evoluția stocurilor materialelor $M_1, M_2, \ldots, M_f$ se modelează prin construcția:

$$y_j = a_1 X_{1j} + a_2 X_{2j} + a_3 X_{3j},\ j = 1,\ 2,\ \ldots,\ f$$

unde:

$f$ - numărul de materiale;

$y_j$ - stocul final al materialului $M_j$;

$X_{1j}$ - stocul inițial al materialului $M_j$;

$X_{2j}$ - intrările prin aprovizionare din materialul $M_j$;

$X_{3j}$ - ieşirile spre consum din materialului $M_j$.

Coeficienții de contribuție au nivelurile $a_1 = 1; a_2 = 1; a_3 = -1$.

În contabilitate, pentru conturile $CON_1, CON_2, \dots, CON_r$, se înregistrează în partea de debit cheltuielile $CD_{ij}$, $i = 1, 2, \dots, r$, $j = 1, 2, \dots, R_i$, iar în partea de credit cheltuielile $CC_{ik}, i = 1, 2, \dots, r, k = 1, 2, \dots, O_i$, unde:

$r$ – numărul de conturi;

$R_i$ – numărul de înregistrări în partea de debit a contului $CON_i$;

$O_i$ – numărul de operații în partea de credit a contului.

Pentru calculul soldului contului debitor $CON_p$, notat $S_p$, se evaluează expresia:

$$S_p = \sum_{j=1}^{R_p} a_j \cdot CD_{pj} + \sum_{o=1}^{O_k} b_o \cdot CC_{po},$$

unde coeficienții de contribuție au valorile: $a_1 = a_2 = \dots = a_{R_p} = 1$, respectiv $b_1 = b_2 = \dots = b_{O_k} = -1$.

**Modelele econometrice** reprezintă o categorie deosebit de importantă de modele utilizate preponderent pentru studierea fenomenelor și proceselor la nivel macroeconomic. Se consideră o mulțime de variabile rezultative $Y_1, Y_2, \dots, Y_M$, și o serie de factori $F_1, F_2, \dots, F_N$, cărora li se asociază respectiv variabilele $X_1, X_2, \dots, X_N$.

Atât pentru variabilele rezultative cât și pentru variabilele exogene se efectuează măsurători pentru $k$ momente, rezultând un număr de MN serii de timp, unde MN=M+N cu $M$ - numărul de variabile rezultative, $N$ - numărul de variabile exogene.

Este important să se stabilească: lungimea seriilor de timp, modul în care se alege momentul $T_1$, uniformizarea intervalelor de culegere a datelor și procedurile de culegere a datelor.

Analiza economică determină stabilirea structurilor modelelor econometrice, rezultând matricea $E$, a dependențelor, dată în tabelul 1.

**Tabelul 1. Matricea dependențelor E**

| | $X_1 \ X_2 \dots X_j \dots X_N$ |
|---|---|
| $Y_1$ | $e_{11} \ e_{12} \dots e_{1j} \dots e_{1N}$ |
| $Y_2$ | $e_{21} \ e_{22} \dots x_{j2} \dots e_{2N}$ |
| $\vdots$ | $\vdots$ |
| $Y_k$ | $e_{k1} \ e_{k2} \dots e_{kj} \dots e_{kN}$ |
| $\vdots$ | $\vdots$ |
| $Y_M$ | $e_{M1} \ e_{M2} .. e_{Mj} \dots e_{MN}$ |

unde $e_{kj} = \begin{cases} 1, \ daca \ exista \ o \ legatura \\ 0, \ in \ rest \end{cases}$.

În cazul modelului econometric Klein-Goldberger:

$$\begin{aligned}
C_t &= \alpha_0 + \alpha_1 P_t + \alpha_3 (W_t^p + W_t^g) + z_1 \\
I_t &= \beta_0 + \beta_1 P_t + \beta_2 P_{t-1} + \beta_3 K_{t-1} + z_2 \\
W_t^p &= \gamma_0 + \gamma_1 X_t + \gamma_2 X_{t-1} + \gamma_3 A_t \\
X_t &= C_t + I_t + G_t \\
P_t &= X_t - T_t - W_t^p \\
K_t &= K_{t-1} + I_t,
\end{aligned}$$

cu semnificația:

- C – consumul;
- I – investiţia;
- Wp – salariile private;
- X – cererea la echilibru;
- P – profiturile private;
- K – nivelul capitalului;
- G – cheltuielile guvernamentale;
- T – taxele indirect plus exportul net;
- Wg – salariile bugetare;
- A – tendinţa măsurată din 1931.

Matricea dependenţelor este:

**Tabelul 2. Matricea dependenţelor Klein-Goldberger**

| | $C$ | $I$ | $W_p$ | $X$ | $P$ | $K$ | $G$ | $T$ | $W_g$ | $A$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $C_t$ | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| $I_t$ | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| $W_t^p$ | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| $X_t$ | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| $P_t$ | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| $K_t$ | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |

Rezultă că se estimează coeficienţii ecuaţiilor modelelor econometrice, pentru variabila rezultativă $Y$ :

$$\hat{y}_i = \sum_{j=1}^{N} a_{ij}\alpha_{ij}x_{ij}$$

unde:

M – numărul de ecuaţii;

$a_{ij}$ – coeficienţi de contribuţie $a_{ij} \in \{-1, 0, 1\}$ ;

$\alpha_{ij}$ – coeficienţii modelului, estimaţi prin metoda celor mai mici pătrate;

$\hat{y}_i$ – nivelul estimat al variabilei rezultative $\hat{y}_i$ din modelul Klein-Goldberger.

**Modele de optimizare liniară** includ în structura lor funcţii de forma:

$$\{\min/\max\} f(x) = \sum_{j=1}^{n} a_j c_j x_j$$

unde:

$n$ – numărul de produse;

$a_{ij}$ – coeficienţii de contribuţie definiţi pe mulţimea $\{-1,0,1\}$ ;

$c_j$ – coeficienţii de multiplicare;

$x_j$ – nivelul variabilelor asociate factorilor, care influenţează procesele de maximizare sau de minimizare.

Restricţiile modelelor liniare au forma:

$$\sum_{j=1}^{m} a_{ij}\alpha_{ij}x_j \le b_i, i = 1,2,\ldots,n$$

unde:

$a_{ij}$ - coeficienţii de contribuţie pentru ecuaţia $i$;

$b_{ij}$ - nivel limitativ pentru definirea unui capăt al intervalului de variabile pentru resursele utilizate;

$\alpha_{ij}$ - consumul unitar de resursă de tip $i$, asociată factorului de influenţă;

$m$ - numărul de restricţii.

**Modelele neliniare** reprezintă o largă categorie utilizată în studierea interdependenţelor dintre factori. Marea varietate de modele neliniare produce efecte variate în colectarea şi dezvoltarea sistematizată a lor, întrucât în descrierea modelelor de acest tip trebuie definite reguli care să conducă la descrieri corecte şi la implementări cu grad de cuprindere deosebit de ridicat.

Sunt situaţii în care modelele neliniare, prin transformări convenabile, sunt transformate în modele liniare.

De exemplu, modelul neliniar:

$$y = A \cdot X^{\alpha} \cdot W^{\beta}$$

prin logaritmare, conduce la structura:

$$\log y = \log A + \alpha \log X + \beta \log Y .$$

Efectuând înlocuirile $y' = \log y$, $A' = \log A$, $X' = \log X$, $W' = \log W$, se obţine modelul liniar:

$$y' = A' + \alpha X' + \beta W'$$

care se rescrie în forma

$$y' = \sum_{i=1}^{n} a_i b_i X_i ,$$

unde $X_1 = A', X_2 = X', X_3 = W'$, $b_1 = 1, b_2 = \alpha, b_3 = \beta$ şi $a_1 = a_2 = a_3 = 1$.

Toate acestea conduc la ideea găsirii unor tehnici şi metode de stocare şi gestionare a diversităţii de modele economice, folosind software adecvat.

## 2. Baza de modele economice

Baza de modele este o construcţie complexă, în care sunt incluse liste cu:

- variabile dependente şi variabile independente;
- structuri de modele liniare şi neliniare;
- seturi de date înregistrate şi seturi de date generate;
- implementări ale algoritmilor de estimare coeficienţi;
- proceduri pentru verificarea de ipoteze;
- proceduri pentru calcul de valori estimate;
- proceduri de ierarhizare a modelelor.

Aceste componente deosebit de importante sunt elemente cu care este populată baza de modele. Pentru ca baza de modele să devină operaţională, trebuie să existe un sistem de gestiune a sa.

În primul rând, sistemul de gestiune trebuie să opereze distinct cu seturile de date, cu procedurile şi cu structurile de modele.

În al doilea rând, funcţiile sistemului de gestiune trebuie să efectueze regăsirea rapidă a seturilor de date, a structurilor de modele şi a procedurilor în vederea asigurării derulării de procese în concordanţă cu cerinţele economistului analist.

În al treilea rând, sistemul de gestiune trebuie să fie înzestrat cu funcţii care să permită adăugarea de seturi de date, adăugarea de modele şi adăugarea de proceduri. Trebuie schimbată optica prin care actualizarea bazei de modele presupunea ştergeri de date/modele/proceduri sau modificarea unor părţi ale acestora cu noi secvenţe.

Acceptarea funcției de actualizare exclusive prin adăugare vine să aducă o concordanță între modul natural de concepere a evoluției, cu reflectarea corespunzătoare a acesteia în plan informatic.

Sistemul de gestiune a bazei de modele operează cu entități neomogene, aspect important în asigurarea consistenței fluxurilor.

În al patrulea rând, se asigură caracter deschis, utilizatorii având la dispoziție posibilitatea de a defini algoritmi proprii pentru interpolare și extrapolare, pentru generare de numere pseudoaleatoare, pentru estimarea de coeficienți, pentru a implementa criterii proprii de selecție a modelelor.

În al cincilea rând, definirea conceptelor specifice privind regăsirea, selecția, extragerea, vizarea tripletelor (date, structuri de modele, proceduri) care grupează prelucrări complexe.

În al șaselea rând, se produce o creștere a gradului de generalizare pentru conceptual de tranzacție care, în cazul bazei de modele, presupune traversarea unor fluxuri în care se operează simultan cu seturi de date, cu structuri de date și cu proceduri.

Noul conglomerat, mai complex decât structura obiectuală ce include operanzi și operatori, dezvoltă o nouă proiecție asupra filosofiei de a proiecta un sistem de gestiune, sistemul de gestiune a bazei de modele, în care sunt incluse, pe lângă prelucrările deja uzuale, noi tipologii specifice proceselor implementate în domeniul inteligenței artificiale.

## 3. Ortogonalitatea modelelor

Se consideră o bază de modele în care sunt deja stocate modelele $M_1, M_2, \dots, M_n$. Se pune problema adăugării unui nou model, $M_{n+1}$. Acest lucru se efectuează dacă și numai dacă modelul $M_{n+1}$ este diferit de modelele deja existente.

**Modelele cu structură identică** sunt acelea care au același număr de variabile, același număr de ecuații, același număr de operatori, iar prin dezvoltarea scrierii poloneze inverse, pozițiile operanzilor și operatorilor sunt aceleași.

Modelele:

$$M_1 : y = a \cdot x + b \cdot z + c$$
$$M_2 : w = d \cdot u + h \cdot v + g$$

sunt cu structură identică, iar modelele:

$$M_3 : y = ax + be^t$$
$$M_4 : u = ce^w + dw$$

sunt identice din punct de vedere structural, adunarea fiind comutativă.

Complexitățile neponderate se calculează după următoarea

$$C(M_i) = n_i \log_2 n_i + m_i \log_2 m_i$$

unde:

- $n_i$ numărul de operanzi ai modelului;
- $m_i$ numărul de operatori ai modelului.

Pentru calculul complexităților ponderate, fiecare operator primește o pondere $w_j$, pentru exemplele prezente aici, ponderile operatorilor $=,+,*,()^2$ fiind respectiv 1, 2, 3, 4. Formula este:

$$C(M_i) = n_i \log_2 n_i + S \log_2 S$$

unde:

$$S = \sum_{j=1}^{m_i} w_j \ .$$

Pentru a vedea dacă două modele sunt identice, se procedează, mai întâi, la calculul complexităţii neponderate, $C(M_1) = C(M_2) = C(M_3) = C(M_4) = 27{,}11$.

La o primă analiză, rezultă că modelele au complexităţi neponderate identice.

Pentru modelele $M_1$ şi $M_2$, complexităţile ponderate sunt date de relaţia: $C(M_1) = C(M_2) = 53{,}56$, ceea ce arată că modelele $M_1$ şi $M_2$ au un grad de asemănare foarte ridicat, mai mult, din punct de vedere structural sunt identice, $C(M_3) = C(M_4) = 63{,}61$.

Modelele $M_3$ şi $M_4$ sunt identice din punct de vedere structural şi diferite de $M_1$ şi $M_2$.

Modelele incluse sunt acelea care, în scrierea poloneză inversă, apar sub formă de subşiruri ale unor şiruri. Pentru modelele:

$$M_5: y = ax + bz + c$$

$$M_6: y = a \cdot x + b \cdot z + d \cdot u + c$$

se observă că $M_6$ se deduce din $M_5$ la care se adaugă termenul $d \cdot u$. Înseamnă că modelul $M_5$ este inclus în modelul $M_6$. Complexităţile sunt: $C(M_5) = 27.11$ şi $C(M_6) = 39.5$.

Raportul de asemănare, $A$, este dat de relaţia:

$$A = \frac{K}{N}$$

unde:

$N$ - numărul maxim de componente;

$K$ - numărul de componente identice.

Pentru modelele $M_5$ şi $M_6$, $N = \max\{11,14\} = 14$, unde valorile din paranteze reprezintă numărul de variabile şi de operatori. Elementele comune se reunesc într-o mulţime $\{y, =, a, \cdot(produs), x, +(adunare), b, \cdot(produs), z, +, c\}$ al cărei cardinal este 11, rezultând A = 0,78.

Complexităţile ponderate sunt $C(M_5) = 53.56$ şi $C(M_6) = 104.71$.

Dacă modelul liniar $M_k$ este inclus în modelul $M_j$, unde

$$M_k:\ y = a_0 + \sum_{k=1}^{n} a_i x_i$$

$$M_j:\ y = a_0 + \sum_{k=1}^{n} a_i x_i + \sum_{i=n+1}^{n+r} a_i x_i$$

complexităţile calculate sunt:

$$C(M_k) = (n+2) \cdot \log_2 (n+2) + (2n) \cdot \log_2 (2n)$$

$$C(M_j) = (n+r+2) \cdot \log_2 (n+r+2) + 2(n+r) \cdot \log_2 2(n+r)$$

**Modelele ortogonale** sunt acelea în care operanzii şi operatorii sunt diferiţi, fără a lua în considerare operatorul de atribuire:

$$M_7:\ y = a \cdot b^t$$

$$M_8:\ y = cx + dz + h$$

Operatorii modelului $M_7$ sunt constituiţi în şirul: * şi ( ), iar operatorii modelului $M_8$ alcătuiesc şirul: *,+,*,+. Este evident că şirurile sunt total diferite.

Operanzii din modelul $M_7$ se constituie în şirul: y, a, b, t, pe când cei din modeleul $M_8$ se regăsesc în: z, c, x, d, z, h. Şirurile sunt, de asemenea, diferite.

Gradul de ortogonalitate se obţine din relaţia:

$O(M_k, M_j) = 1 - A(M_k, M_j)$.

Dacă indicatorul $O(M_k, M_j) = 1$, rezultă că modelele sunt ortogonale.

În cazul în care în baza de modele ce conţine modelele $M_1, M_2, \ldots, M_n$ şi se doreşte adăugarea modelului $M_{n+1}$, dacă $O(M_i, M_{n+1}) = 1, \forall i = 1,2,..,n$, modelul $M_{n+1}$ este într-adevăr un model nou şi, prin includerea lui în baza de modele, aceasta sporeşte din punct de vedere calitativ.

**Modelele de aceeaşi clasă** sunt modelele care nu diferă din punct de vedere structural. Modelele:

$$M_8:\ y = ax + bz + cw + d$$
$$M_9:\ y = ax + d$$

sunt modele liniare, provin din aceeaşi clasă pentru că includ în alcătuirea lor: coeficienţi, variabile, operatori de adunare, operatori de înmulţire şi termeni de acelaşi grad egal cu 1.

Se calculează complexitatea:

$C = n_1 \log_2 n_1 + n_2 \log_2 n_2 + n_3 \log_2 n_3$

unde:

$n_1$ - numărul de termeni diferiţi;

$n_2$ - numărul de operanzi diferiţi;

$n_3$ - numărul de grade diferite.

În mod corespunzător, complexităţile modelelor $M_8$ şi $M_9$ sunt $C(M_8) = C(M_9) = 2\log_2 2 + 3\log_2 3 + 2\log_2 2 = 8.75$.

Modelele:

$$M_{10}:\ y = ax_1^2 + bx_2^2 + cx_3^2 + d + ex_1 + fx_2 + gx_3$$
$$M_{11}:\ y = ax_1^2 + bx_2^2 + cx_3^2 + dx_4^2 + ex_5^2 + gx_6^2 + hx_7^2 + i + jx_1 + kx_2 + lx_3$$

sunt definite prin:

- coeficienţii modelului $M_{10}: a, b, c, d, e, f, g$.
- coeficienţii modelului $M_{11}:\ a, b, c, d, e, g, h, i, j, k, l$.
- variabilele modelului $M_{10}:\ y, x_1, x_2, x_3$.
- variabilele modelului $M_{11}:\ y, x_1, x_2, x_3, x_4, x_5, x_6, x_7$.
- operatorii diferiţi ai modelului $M_{10}:\ =, +, ()^2, *$.
- operatorii diferiţi ai modelului $M_{11}$ model are: $=, +, *, ()^2$.
- gradele termenilor modelelor $M_{10}$ şi $M_{11}$: 0,1,2.

Se calculează complexităţile modelelor, obţinându-se $C(M_{10}) = C(M_{11}) = 3\log 3 + 4\log 4 + 3\log 3 = 175$.

Rezultă că $M_{10}$ şi $M_{11}$ aparţin aceleiaşi clase.

**Modelele asemenea** sunt acelea care conţin structuri dominante. Modelele:

$$M_{12}: \ y = ax_1 + bx_2x_3 + ce^{x_4} + dx_5 \cdot \log(x_6) + g$$

$$M_{13}: \ y = ax_1^2 + b\frac{x_2}{x_3} + ce^{x_4} + dx_5 \cdot \log(x_6) + g$$

au în structura lor termenii comuni, reuniți în subexpresia $ce^{x_4} + dx_5 \cdot \log(x_6) + g$. Complexitățile neponderate ale celor două modele sunt: $C(M_{12}) = 86.03$, respectiv $C(M_{13}) = 96.21$.

Complexitatea subexpresiei comune este:

$C(common) = 6\log_2 6 + 7\log_2 7 = 35.16$.

Ponderea complexității subexpresiei comune a două modele $M_i$ și $M_j$ se calculează astfel:

$$\rho_i = \frac{C(M_i \cap M_j)}{C(M_i)}$$

$$\rho_j = \frac{C(M_i \cap M_j)}{C(M_j)}$$

Ponderea complexităţi subexpresiei comune în totalul complexității lui $M_{12}$ este :

$$\rho_{12} = \frac{6\log_2 6 + 7\log_2 7}{12\log_2 12 + 12\log_2 12} = 0.4$$

şi ponderea complexităţi subexpresiei comune în totalul complexității lui $M_{13}$ este:

$$\rho_{13} = \frac{6\log_2 6 + 7\log_2 7}{13\log_2 13 + 13\log_2 13} = 0.36.$$

Dacă $\rho \in [0,0.82]$, modelele au în comun elemente nesemnificative.

Dacă $\rho \in [0.82,0.92]$, modelele au în comun elemente semnificative.

Dacă $\rho \in [0.92,1)$, modelele au un grad foarte ridicat de asemănare.

Pentru $\rho_i = 1$, un anumit model, $M_i$ este inclus în alt model, $M_j$, cu $\rho_j > 0$. Dacă $\rho_i = \rho_j = 1$, modelele sunt identice.

În concluzie, realizarea analizei ortogonalității modelelor economice permite includerea acestora în clase de ortogonalitate a modelelor.

## 4. Ortogonalitatea seturilor de date

Modelele economice vizează laturile calitative şi laturile cantitative ale evoluției unor fenomene sau ale definirii de structuri, în vederea obținerii de informații privind efectele apariției unor modificări în timp.

Orice model economic presupune înregistrarea de date pentru a vedea nivelurile caracteristicilor cu care sunt descrise dinamica unui fenomen şi factorii care o determină.

Setul de date, asociat unei probleme de dinamică, este dat sub forma unui tabel ce include:

- coloană în care se specifică momentele de timp în care au fost efectuate măsurătorile;
- coloană de date reprezentând nivelul măsurat pentru variabila rezultativă;
- câte o coloană de date reprezentând nivelurile măsurate ale fiecărui factor de influență.

Pentru a realiza o prelucrare de calitate trebuie ca seriile incluse în tabel să aibă acelaşi număr de termeni şi măsurarea lor să se realizeze exact la momentul de timp, indicat în tabloul construit conform structurii date, în tabelul 3.

**Tabelul 3. Structură set de date specific proceselor dinamice**

| Moment de timp T | Variabilă rezultativă y | Factor de influență $X_1$ | ... | Factor de influență $X_j$ | ... | Factor de influență $X_m$ |
|---|---|---|---|---|---|---|
| $T_1$ | $y_1$ | $X_{11}$ | ... | $X_{j1}$ | ... | $X_{m1}$ |
| $T_2$ | $y_2$ | $X_{12}$ | ... | $X_{j2}$ | ... | $X_{m2}$ |
| $T_3$ | $y_3$ | $X_{13}$ | ... | $X_{j3}$ | ... | $X_{m3}$ |
| ... | ... | ... | ... | ... | ... | ... |
| $T_k$ | $y_k$ | $X_{1k}$ | ... | $X_{jk}$ | ... | $X_{mk}$ |
| ... | ... | ... | ... | ... | ... | ... |
| $T_n$ | $y_n$ | $X_{1n}$ | ... | $X_{jn}$ | ... | $X_{mn}$ |

unde:

$T_k$ – momentul de timp la care a avut loc înregistrarea valorilor;

$y_k$ – valoarea variabilei rezultative la momentul $T_k$;

$X_{jk}$ – valoarea factorului de influență j la momentul de măsurare $T_k$.

Dacă diferențele

$$T_2 - T_1 = T_3 - T_2 = ... = T_{k+1} - T_k = ... = T_n - T_{n-1}$$

rezultă că înregistrările de date au fost efectuate la intervale egale de timp.

Pe seturile de date se definesc care afectează atât structura, cât şi conținutul acestora, precum:

- adăugarea de noi linii, respectiv efectuarea de măsurători pentru momentele $T_{n+1}$, $T_{n+2}, ...$
- adăugarea de vechi măsurători efectuate la momente ce au precedat momentul $T_1$, respectiv $T_0$, $T_{-1}$, $T_{-2}, ...$
- eliminarea de coloane din tabel, în cazul în care variabilele nu se dovedesc a fi esențiale pentru proces de date, pentru alte momente de timp şi înregistrarea de noi niveluri pentru toate variabilele, incluzând, deci, şi noile variabile adăugate pentru a obține un nou set de date.

Al doilea tip de seturi de date vizează o colectivitate A formată din elementele $a_1$, $a_2$, ..., $a_n$, pentru care se înregistrează niveluri ale caracteristicilor $C_1$, $C_2$, ..., $C_m$, a cărui structură este prezentată în tabelul 4:

**Tabelul 4. Set de date pentru descrierea colectivității A**

| Element | Caracteristica $C_1$ | ... | Caracteristica $C_j$ | ... | Caracteristica $C_m$ |
|---|---|---|---|---|---|
| $a_1$ | $C_{11}$ | ... | $C_{1j}$ | ... | $C_{1m}$ |
| $a_2$ | $C_{21}$ | ... | $C_{2j}$ | ... | $C_{2m}$ |
| ... | ... | ... | ... | ... | ... |
| $a_i$ | $C_{i1}$ | ... | $C_{ij}$ | ... | $C_{im}$ |
| ... | ... | ... | ... | ... | ... |
| $a_n$ | $C_{n1}$ | ... | $C_{nj}$ | ... | $C_{nm}$ |

unde: $a_i$ – elementul i al colectivității A şi $C_{ji}$ – valoarea caracteristicii j pentru elementul $a_i$ al colectivității A.

Operațiile pe acest tip de set de date sunt:

- adăugarea unor noi elemente $a_{n+1}$, $a_{n+2}$, ...
- adăugarea de noi caracteristici $C_{m+1}$, $C_{m+2}$, ... cu reluarea măsurătorilor;
- eliminarea unor elemente din mulțime;
- eliminarea unor caracteristici, de regulă cele ale căror niveluri sunt identice pentru toate elementele colectivității.

Ca şi în cazul primului tip de set de date, pentru elementele colectivității A trebuie definite şi utilizate riguros proceduri pentru măsurarea ortogonalității.

Existenţa mai multor seturi de date ridică problema ortogonalităţii acestora atât sub aspectul structurii, cât şi în ceea ce priveşte conţinutul lor.

Fie seturile de date $S_i$ şi $S_j$. Cele două seturi de date se caracterizează printr-o structură şi un conţinut.

Structura setului de date este dată de prima linie a tabelului cu ajutorul căruia setul de date se reprezintă. Conţinutul setului de date constă în liniile care urmează celei de descriere a structurii.

Analiza ortogonalităţii seturilor de date se desfăşoară pe mai multe niveluri. Continuarea procesului de rafinare a seturilor de date pe elemente structurale, specifice acestora, este condiţionată de identitatea perfectă a seturilor de date, într-un anumit stadiu al analizei.

Fie seturile de date $S_1$ şi $S_2$ care conţin valori privind nivelurile cheltuielilor şi veniturilor a două societăţi comerciale $SC_1$ şi $SC_2$:

**Tabelul 5. Setul de date $S_1$**

| Luna | Cheltuieli (mil. lei) | Venituri (mil. lei) |
|---|---|---|
| Ianuarie | 250 | 300 |
| Februarie | 400 | 350 |
| Martie | 350 | 350 |
| Aprilie | 550 | 450 |
| Mai | 200 | 300 |
| Iunie | 300 | 400 |

respectiv:

**Tabelul 6. Setul de date $S_2$**

| Luna | Cheltuieli (mil. lei) | Venituri (mil. lei) |
|---|---|---|
| Martie | 350 | 350 |
| Iunie | 300 | 400 |
| Ianuarie | 250 | 300 |
| Mai | 200 | 300 |
| Februarie | 400 | 350 |
| Aprilie | 550 | 450 |

Primul nivel al analizei ortogonalităţii seturilor de date $S_1$ şi $S_2$ priveşte structura acestora. Seturile de date sunt perfect identice deoarece au acelaşi număr de linii şi de coloane:

$$nl_1 = nl_2 = 7$$

$$nc_1 = nc_2 = 3$$

unde:

$nl_1$, $nl_2$ – numărul de linii ale seturilor de date $S_1$, respectiv $S_2$;

$nc_1$, $nc_2$ – numărul de coloane ale seturilor de date $S_1$, respectiv $S_2$.

Deoarece identitatea între cele două seturi de date este identică, se trece la un nou nivel de rafinare a acestora.

În continuare, se determină valorile maxime şi minime a valorilor din cele două tabele. Astfel:

$$MIN(S_1) = MIN\ (S_2) = 200$$

$$MAX(S_1) = MAX(S_2) = 550$$

seturile de date fiind identice conform acestui criteriu de asemănare.

Următorul criteriu în analiză constă în determinarea valorilor minime şi maxime pe fiecare linie a celor două seturi de date. Procedând la sortarea crescătoare a valorilor, se obţine următorul tabel:

**Tabelul 7. Valorile ordonate de minim şi maxim pe linii a seturilor de date**

| $Min_1$ | $Max_1$ | $Min_2$ | $Max_2$ |
|---|---|---|---|
| 200 | 300 | 200 | 300 |
| 250 | 300 | 250 | 300 |
| 300 | 350 | 300 | 350 |
| 350 | 400 | 350 | 400 |
| 350 | 400 | 350 | 400 |
| 450 | 550 | 450 | 550 |

Analizând datele prezentate în tabelul de mai sus, se observă că seturile de date $S_1$ şi $S_2$ sunt perfect identice pentru criteriul luat în considerare. Ca urmare, are loc adâncirea procesului de analiză a ortogonalităţii.

Următorul criteriu constă în aplicarea unui minim sau maxim pe fiecare coloană de valori, iar pentru şirul de valori obţinut pentru un set de date se aplică un obiectiv de sens contrar. De exemplu, pentru seturile de date considerate, se determină minimul pe fiecare coloană. Şirurile rezultate sunt:

$S_{min1} = (200, 300)$

$S_{min2} = (200, 300)$

Pentru cele două şiruri se aplică funcţia obiectiv de sens opus:

$max_{S1} = max_{S2} = max\ (200, 300)$

ceea ce înseamnă că seturile de date sunt perfect identice şi din acest punct de vedere.

Ultimul nivel al analizei constă în compararea valorilor în mod individual pentru cele două seturi de date, prin stabilirea de relaţii de ordine, a proporţionalităţii valorilor etc.

În exemplul de faţă, cele două seturi sunt perfect identice, orice criteriu ar fi luat în considerare. În general, un set de date este identificat în mod unic, pe baza unei amprente a setului de date.

Amprenta setului de date cuprinde următoarele elemente:

- structura setului de date exprimată în număr de linii şi coloane;
- valoarea minimă, respectiv maximă a valorilor conţinute în tabel;
- valorile corespunzătoare criteriilor *min(max)*, respectiv *max(min)*;
- valoarea variabilei endogene, profit brut în cazul exemplului considerat.

Seturile de date, utilizate în elaborarea de modele şi în studierea proceselor economice, trebuie să îndeplinească o serie de caracteristici de calitate. Acestea trebuie identificate, măsurate, influenţând întreaga abordare a modelării procesului ca practică şi bază a unor dezvoltări coerente.

În acest context, ortogonalitatea seturilor de date joacă un rol important în dezvoltarea unor modele cu un nivel ridicat al calităţii.

## 5. Concluzii

Bazele de modele includ:

- modele ortogonale;
- modele generative pentru o clasă, de exemplu:

$$y = a_0 + \sum_{i=1}^{n} a_i x_i\ ,\ y = a_0 + \sum_{i=1}^{n} a_i x_1^2\ ,\ \text{etc.}$$

- modele cu subexpresii comune cu pondere limitată;
- modele cu un nivel de includere sub o limită impusă.

Limitele sunt impuse atunci când se configurează baza de modele, pentru a se gestiona riguros varietăţile de modele.

Ortogonalitatea trebuie surprinsă şi în ceea ce priveşte seturile de date. Acestea joacă un rol important în testarea şi aplicarea modelelor în realitate, contribuind la fundamentarea unor decizii corecte.

## Bibliografie


1. **IVAN, I., A. VIŞOIU**: Generator de modele liniare cu argument întârziat. În: Revista de Comerţ, nr. 1, 2004, pp. 47 – 50.
2. **IVAN, I., R. ENYEDI, M. POPA**: Software Cloning by Translation Processes. În: Lucrările Conferinţei Internaţionale „Cibernetica la nivel macroeconomic şi microeconomic", Bucureşti, 2004.
3. **IVAN, I.**: Baze de modele pentru managementul calităţii software, Contract de cercetare, ASE, 2003.
4. **IVAN, I., M. POPA, S. CAPISIZU, B. LUKACS, B. FLORESCU**: Clonarea informatică, Editura ASE, Bucureşti, 2003.
5. **IVAN, I., M. POPA, S. CAPISIZU, B. FLORESCU, L. IVAN**: Characteristics of the Informatics Clones, Master of International Business Informatics Handbook, Editura ASE, Bucureşti, 2003, pp. 207 – 223.
6. **IVAN, I., P. POCATILU, M. POPA, T. MIHAI, L. IVAN**: Data Orthogonality, Master of International Business Informatics Handbook, Editura ASE, Bucureşti, 2003, pp. 235 – 249.
7. **NICA, V. T., V. MĂRĂCINE**: Modelarea firmei, Editura ASE, Bucureşti, 2002.
8. **PECICAN, E.**: Econometrie, Editura ALL, Bucureşti, 1994.
9. **POPA, M.**: Software pentru măsurarea gradului de asemănare a fişierelor arhivate, ASE, Bucureşti, 2002