

Metoda Topsis i GDM

Marcin Szymkowiak

Metody porządkowania liniowego

- Metody porządkowania liniowego – ujęcie teoretyczne
- Metoda TOPSIS - The **T**echnique for **O**rders of **P**reference by **S**imilarity to **I**deal **S**olution
- Metoda GDM - **G**eneralised **D**istance **M**easure

Metody porządkowania liniowego – ujęcie teoretyczne

- Zadaniem metod porządkowania liniowego zbioru obiektów jest uszeregowanie, czyli ustalenie kolejności obiektów lub zbiorów według określonego kryterium.
- Metody porządkowania liniowego stosowane są wówczas gdy można przyjąć pewne nadrzędne kryterium, ze względu na które będzie można uszeregować obiekty od „najlepszego” do „najgorszego”.
- Narzędziem metod porządkowania liniowego jest syntetyczny miernik rozwoju (SMR), będący pewną funkcją agregującą informacje cząstkowe zawarte w poszczególnych zmiennych i wyznaczoną dla każdego obiektu ze zbioru obiektów A.

Metody porządkowania liniowego – ujęcie teoretyczne

- Dany jest co najmniej dwuelementowy i skończony zbiór obiektów $A = \{A_i\}_{i=1}^n = \{A_1, \dots, A_n\}$.
- Istnieje pewne nadrzędne syntetyczne kryterium porządkowania elementów zbioru A , które nie podlega pomiarowi bezpośredniemu (na przykład poziom rozwoju badanego produktu na tle produktów konkurencyjnych, ocena wstępnie wyselekcjonowanych koncepcji produktu).
- Dany jest skończony zbiór zmiennych, merytorycznie związany z syntetycznym kryterium porządkowania; zmienne mają charakter preferencyjny, tzn. wyróżnia się wśród nich stymulanty, destymulanty i nominanty.
- Zmienne służące do opisu obiektów są mierzone przynajmniej na skali porządkowej. Jeżeli zmienne opisujące obiekty mierzone są na skali przedziałowej lub ilorazowej, należy sprowadzić je do porównywalności poprzez normalizację.
- Relacją porządkującą elementy zbioru A jest relacja większości lub mniejszości, dotycząca ilczbowych wartości syntetycznego miernika rozwoju.

Metody porządkowania liniowego – ujęcie teoretyczne

- Sformułowanie celu analizy oraz wstępnych hipotez badawczych.
- Dobór cech.
- Zebranie materiału statystycznego.
- Wybór metody porządkowania liniowego.
- Utworzenie rankingu obiektów za pomocą wybranej metody.
- Weryfikacja wyników porządkowania.
- Interpretacja wyników porządkowania.

Metody porządkowania liniowego – ujęcie teoretyczne

- Cecha diagnostyczna X_k jest **stymulantą** (S), gdy dla każdych dwóch jej wartości x_{ik}^S, x_{jk}^S odnoszących się do obiektów A_i, A_j , spełniony jest warunek $x_{ik}^S > x_{jk}^S \Rightarrow A_i \succ A_j$ (\succ oznacza dominację obiektu A_i nad obiektem A_j) Innymi słowy cecha X_k jest stymulantą, gdy jej wyższe wartości są pożądane.
- Cecha diagnostyczna X_k jest **destymulantą** (D), gdy dla każdych dwóch jej wartości x_{ik}^D, x_{jk}^D odnoszących się do obiektów A_i, A_j , spełniony jest warunek $x_{ik}^D > x_{jk}^D \Rightarrow A_i \prec A_j$ (\prec oznacza dominację obiektu A_j nad obiektem A_i) Innymi słowy cecha X_k jest destymulantą, gdy jej niższe wartości są pożądane.
- Cecha diagnostyczna X_k jest **nominantą** (N), gdy dla każdych dwóch jej wartości x_{ik}^N, x_{jk}^N odnoszących się do obiektów A_i, A_j :
 - jeżeli $x_{ik}^N, x_{jk}^N \leq nom_k$ to $x_{ik}^N > x_{jk}^N \Rightarrow A_i \succ A_j$,
 - jeżeli $x_{ik}^N, x_{jk}^N > nom_k$ to $x_{ik}^N > x_{jk}^N \Rightarrow A_i \prec A_j$.

Najbardziej pożądaną z punktu widzenia rozpatrywanego obiektu lub zjawiska jest wartość nominalna, natomiast wartości mniejsze bądź też większe od wartości nominalnej są mniej korzystne.

Metody porządkowania liniowego – ujęcie teoretyczne

Konkretna technika normalizacji zmiennych zależy od ich charakteru, tj. czy są one stymulantami, destymulantami czy nominantami. Rzeczą niezmiennie istotną, jest również to, aby stosowane procedury normalizacyjne spełniały następujące wymogi:

- pozbawienie mian (jednostek), w których wyrażone są cechy diagnostyczne,
- sprowadzenie rzędu wielkości cech diagnostycznych do porównywalności, co oznacza wyrównanie zakresów zmienności cech, a w konsekwencji możliwość ich dodawania,
- równość długości przedziałów zmienności wartości wszystkich cech unormowanych,
- możliwość normowania cech diagnostycznych przyjmujących wartości zarówno dodatnie jak i ujemne,
- możliwość normowania cech przyjmujących wartość równą zeru,
- nieujemność wartości cech unormowanych,
- istnienie prostych formuł – w ramach danej procedury normalizacyjnej – ujednolicających charakter zmiennych.

Metoda Topsis

- **Metoda Topsis** zaliczana jest do tzw. wzorcowych metod porządkowania liniowego obiektów opisywanych za pomocą wielu cech.
- Została ona opracowana przez Hwanga i Yoon (1981) a następnie rozwijana przez Yoon (1987), Hwanga, Lai i Liu (1993).
- Idea metody TOPSIS polega na określeniu odległości rozpatrywanych obiektów od rozwiązania idealnego i antyidealnego.
- Kończącym rezultatem analizy jest wskaźnik syntetyczny tworzący ranking badanych obiektów.
- Za najlepszy obiekt uważa się ten, który ma najmniejszą odległość od rozwiązania idealnego i jednocześnie największą od rozwiązania antyidealnego.

Procedura postępowania w metodzie TOPSIS

- 1 Utworzenie znormalizowanej macierzy danych według formuły:

$$z_{ij} = \frac{x_{ij}}{\sqrt{\sum_{i=1}^n x_{ij}^2}}, \quad (1)$$

gdzie $i = 1, \dots, n$ jest liczbą porządkowanych obiektów a $j = 1, \dots, k$ liczbą opisujących je cech.

- 2 Uwzględnienie wag przypisanych poszczególnym cechom:

$$v_{ij} = w_j z_{ij}. \quad (2)$$

Procedura postępowania w metodzie TOPSIS

- 3 Ustalenie wektora wartości rozwiązania idealnego a^+ i antyidealnego a^- :

$$a^+ = (a_1^+, \dots, a_k^+) = \left\{ \left(\max_{i=1, \dots, k} v_{ij} \mid j \in J_Q \right), \left(\min_{i=1, \dots, k} v_{ij} \mid j \in J_C \right) \right\}, \quad (3)$$

$$a^- = (a_1^-, \dots, a_k^-) = \left\{ \left(\min_{i=1, \dots, k} v_{ij} \mid j \in J_Q \right), \left(\max_{i=1, \dots, k} v_{ij} \mid j \in J_C \right) \right\}, \quad (4)$$

gdzie J_Q to zbiór stymulant, J_C to zbiór destymulant.

Procedura postępowania w metodzie TOPSIS

- 4 Obliczenie odległości euklidesowych badanych obiektów od rozwiązania idealnego i antyidealnego:

$$d_i^+ = \sqrt{\sum_{j=1}^k (v_{ij} - a_j^+)^2}, \quad \text{dla } i = 1, \dots, n \quad \text{oraz } j = 1, \dots, k, \quad (5)$$

$$d_i^- = \sqrt{\sum_{j=1}^k (v_{ij} - a_j^-)^2}, \quad \text{dla } i = 1, \dots, n \quad \text{oraz } j = 1, \dots, k. \quad (6)$$

- 5 Wyznaczenie współczynnika rankingowego określającego podobieństwo obiektów do rozwiązania idealnego:

$$R_i = \frac{d_i^-}{d_i^+ + d_i^-}. \quad (7)$$

- 6 Największa wartość współczynnika R_i wskazuje na rozwiązanie (obiekt) najlepsze w rozpatrywanym problemie porządkowania liniowego.
- 7 Metoda TOPSIS jest zaimplementowana w pakiecie *topsis* programu R.

Metoda GDM

- W metody wzorcowych tworzy się tak zwany obiekt wzorcowy czyli obiekt o pożądanym wartościach zmiennych wejściowych. Miara syntetyczna powstaje na podstawie pomiaru odległości pomiędzy obserwowanym obiektem, a obiektem wzorcowym.
- Jedną z najnowszych metod wzorcowych wykorzystywanych w porządkowaniu liniowym jest uogólniona miara odległości (GDM). Metoda ta opiera się na obliczaniu miary syntetycznej opartej na badaniu odległości od wzorca.
- Zmienne wejściowe podlegają normalizacji, a następnie stymulacji. Wówczas można już obliczyć współrzędne obiektu wzorcowego, na przykład korzystając ze wzoru:

$$z_{0j} = \max_i \{z_{ij}\}, \text{ dla } z_j^S \quad (8)$$

gdzie: z_j^S - znormalizowana wartość j -tej cechy będącej stymulantą.

Uogólniona miara odległości GDM

- Wartość wskaźnika syntetycznego w uogólnionej mierze odległości przyjmuje postać:

$$d_{ik} = \frac{1}{2} - \frac{\sum_{j=1}^m (z_{ij} - z_{kj})(z_{kj} - z_{ij}) + \sum_{j=1}^m \sum_{\substack{l=1 \\ l \neq i, k}}^n (z_{ij} - z_{lj})(z_{kj} - z_{lj})}{2 \left[\sum_{j=1}^m \sum_{l=1}^n (z_{ij} - z_{lj})^2 \cdot \sum_{j=1}^m \sum_{l=1}^n (z_{kj} - z_{lj})^2 \right]^{\frac{1}{2}}} \quad (9)$$

gdzie: z_{ij} (z_{kj} , z_{lj}) - i-ta (k-ta, l-ta) obserwacja na j-tej zmiennej.

- Jest to ogólna postać miary i jej wynik to odległość obiektu i-tego od obiektu k-tego.

Uogólniona miara odległości GDM

- Na potrzeby zastosowania tej miary do obliczenia odległości od wzorca należy tą formułę nieznacznie zmodyfikować. Kolejny wzór wyznacza wartość wskaźnika syntetycznego, który jest odległością i-tego obiektu od obiektu wzorcowego.




$$d_{i0} = \frac{1}{2} - \frac{\sum_{j=1}^m (z_{ij} - z_{0j})(z_{0j} - z_{ij}) + \sum_{j=1}^m \sum_{\substack{l=1 \\ l \neq i}}^n (z_{ij} - z_{lj})(z_{0j} - z_{lj})}{2 \left[\sum_{j=1}^m \sum_{l=1}^n (z_{ij} - z_{lj})^2 \cdot \sum_{j=1}^m \sum_{l=1}^n (z_{0j} - z_{lj})^2 \right]^{\frac{1}{2}}} \quad (10)$$

- Wartości wskaźnika syntetycznego zawierają się w przedziale [0;1] przy czym wyższe wartości są mniej pożądane, ponieważ oznacza to, że obiekt leży dalej od wzorca.
- Metoda GDM jest zaimplementowana w pakiecie *clusterSim* programu R.

Literatura

- Gatnar E., Walesiak M. (red.) (2004), *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*, Wydawnictwo Akademii Ekonomicznej im. Oskara Langego we Wrocławiu, Wrocław 2004.
- Hwang C.L., Yoon K. (1981), *Multiple Attribute Decision Making: Methods and Applications*, New York: Springer-Verlag.
- Hwang C.L., Lai Y.J., Liu T.Y. (1993), *A new approach for multiple objective decision making*, Computers and Operational Research, 20: 889–899.
- Yoon K. (1987), *A reconciliation among discrete compromise situations*, Journal of Operational Research Society, 38. pp. 277–286.
- Walesiak M. (1996), *Metody analizy danych marketingowych*, Wydawnictwo Naukowe PWN, Warszawa.

Dane kontaktowe

 Marcin Szymkowiak
 m.szymkowiak@ue.poznan.pl
 +48 609 689 678