175/2011

Raport Badawczy

Research Report

RB/70/2011

Efektywne metody wyznaczania mediany Kemeny'ego

H. Bury, D. Wagner

Instytut Badań Systemowych
Polska Akademia Nauk

Systems Research Institute
Polish Academy of Sciences

POLSKA AKADEMIA NAUK

Instytut Badań Systemowych

ul. Newelska 6

01-447 Warszawa

tel.: (+48) (22) 3810100

fax: (+48) (22) 3810105

Kierownik Zakładu zgłaszający pracę:
Prof. zw. dr hab. inż. Janusz Kacprzyk

Warszawa 2011

Efektywne metody wyznaczania mediany Kemeny'ego

Wprowadzenie

Istnieje wiele sposobów upraszczania zadania wyznaczania mediany Kemeny'ego, które w przypadku ogólnym stanowi problem NP-trudny. W pracy przedstawiono jedną z takich metod.

Założymy, że dany jest zbiór obiektów $\mathcal{O} = \{O_1, \ldots, O_n\}$ oraz zbiór ekspertów $\mathcal{K} = \{1, ..., K\}$. Zadaniem ekspertów jest uporządkowanie zbioru obiektów zgodnie z przyjętym kryterium lub zbiorem kryteriów. Zakładamy również, że wszystkie obiekty ze zbioru $\mathcal{O}$ są porównywane ze sobą oraz, że w opiniach ekspertów mogą występować obiekty równoważne. W dotychczasowych rozważaniach posługiwano się przede wszystkim macierzami porównań parami oraz wyznaczaną na ich podstawie tzw. macierzą strat R. Celem pracy jest zwrócenie uwagi na fakt, że można uzyskać interesujące wyniki wprowadzając do rozważań macierz rozkładu głosów ekspertów L.

Zaproponowana metoda wykorzystuje twierdzenie Litvaka [3] podające warunek konieczny i dostateczny, który musi spełniać uporządkowanie, aby stanowiło medianę. Można pokazać – wykorzystując to twierdzenie oraz zależności między macierzami R i L – że wyznaczanie mediany sprowadza się do wyznaczenia relacji większości, jaką powinny spełniać obiekty tworzące szukane uporządkowanie. Następnie należy dokonać eliminacji ze zbioru wszystkich możliwych uporządkowań n obiektów tych uporządkowań, w których wymagane relacje większości między obiektami nie są spełnione. Jeżeli w wyniku tego postępowania otrzymano więcej niż jedno uporządkowanie, należy wybrać to, którego odległość (w sensie Kemeny'ego) od zbioru uporządkowań podanych przez ekspertów jest najmniejsza.

1. Definicje mediany Kemeny'ego. Podejście klasyczne [1]

Oceny pary obiektów podane przez ekspertów mogą mieć postać (dla uproszczenia zapisu pominięto indeks eksperta):

$O_i \succ O_j$, co oznacza, że obiekt O_i jest – ze względu na przyjęte kryterium (zbiór kryteriów) – lepszy od obiektu O_j,

$O_i \approx O_j$, co oznacza, że obiekty O_i oraz O_j są – ze względu na przyjęte kryterium (zbiór kryteriów) – równoważne,

$O_i \prec O_j$, co oznacza, że obiekt O_i jest – ze względu na przyjęte kryterium (zbiór kryteriów) – gorszy od obiektu O_j.

Stosowany jest również zapis w postaci uporządkowania

$$P = \{O_{i_1}, O_{i_2}, \ldots, (O_{i_s}, O_{i_t}), \ldots, O_{i_{n-1}}, O_{i_n}\}, \tag{1}$$

który oznacza, że obiekt O_{i_1} jest lepszy od obiektu O_{i_2}; obiekty równoważne (O_{i_s}, O_{i_t}) są ujęte w nawiasach.

Dla ocen par obiektów, które podaje ekspert o numerze k, tworzymy macierz porównań obiektów parami, której elementy są określone następująco

$$A^k = [a_{ij}^k], \text{gdzie } a_{ij}^k = \begin{cases} 1 & \text{jeżeli } O_i \succ^k O_j \\ 0 & \text{jeżeli } O_i \approx^k O_j, \\ -1 & \text{jeżeli } O_i \prec^k O_j \end{cases} \quad a_{ii} = 0, i = 1, ..., n. \tag{2}$$

W dalszych rozważaniach przyjmujemy, że eksperci podają swoje oceny w postaci uporządkowań.[1)]

Zakładamy, że mamy uporządkowanie P^k podane przez eksperta o numerze k oraz dane uporządkowanie P.

Definicja 1 [1].

Odległość między parą obiektów O_i, O_j w uporządkowaniu P oraz parą tych samych obiektów w uporządkowaniu P^k jest definiowana następująco:

$$d_{ij}(P, P^k) = \left| a_{ij} - a_{ij}^k \right|. \tag{3}$$

Definicja 2 [1].

Odległość między uporządkowaniami P oraz P^k jest definiowana następująco:

$$d(P, P^k) = \sum_{i=1}^{n-1} \sum_{j>i}^{n} d_{ij}(P, P^k) = \sum_{i=1}^{n-1} \sum_{j>i}^{n} \left| a_{ij} - a_{ij}^k \right|. \tag{4}$$

Przyjmiemy, że w uporządkowaniu P obiekt O_i poprzedza O_j, czyli $a_{ij}=1$. Wprowadzamy oznaczenie

$$r_{ij}^k = \left| a_{ij}^k - a_{ij} \right| = \left| a_{ij}^k - 1 \right| \qquad k = 1, ..., K. \tag{5}$$

Współczynniki r_{ij}^k mogą przyjmować następujące wartości

[1)] Mając dane uporządkowanie można na jego podstawie uzyskać – w sposób jednoznaczny – macierz porównań obiektów parami. Należy podkreślić, że znajomość tej macierzy nie zawsze umożliwia wyznaczenie odpowiadającego jej uporządkowania. Taka sytuacja ma miejsce, gdy podane oceny par obiektów nie są przechodnie.

$$r_{ij}^k = |a_{ij}^k - 1| = \begin{cases} 0 & \text{jeżeli } a_{ij}^k = 1 \\ 1 & \text{jeżeli } a_{ij}^k = 0 \\ 2 & \text{jeżeli } a_{ij}^k = -1 \end{cases}. \tag{6}$$

Wzór (4) przybiera zatem postać

$$d(P, P^k) = \sum_{i=1}^{n-1} \sum_{j>i}^{n} |a_{ij} - a_{ij}^k| = \sum_{i=1}^{n-1} \sum_{j>i}^{n} |a_{ij}^k - 1| = \sum_{i=1}^{n-1} \sum_{j>i}^{n} r_{ij}^k . \tag{7}$$

Wprowadzamy oznaczenie $r_{ij} = \sum_{k=1}^{K} r_{ij}^k$. (8)

Definicja 3 [3].

Odległość $d(P, P^{(k)})$ uporządkowania P od zbioru uporządkowań

$P^{(k)} = \{P^1, ..., P^K\}$ podanych przez ekspertów wynosi

$$d(P, P^{(k)}) = \sum_{k=1}^{K} d(P, P^k) = \sum_{k=1}^{K} \sum_{i=1}^{n-1} \sum_{j>i}^{n} r_{ij}^k = \sum_{i=1}^{n-1} \sum_{j>i}^{n} r_{ij} . \tag{9}$$

Macierz $R=[r_{ij}]$ nosi nazwę macierzy strat. Ze wzoru (6) wynika, że wartości elementów macierzy R zależą wyłącznie od postaci uporządkowań podanych przez ekspertów.

Definicja 4 [3].

Medianą Kemeny'ego nazywamy takie uporządkowanie $\widetilde{P}$, że

$$d(\widetilde{P}, P^{(k)}) = \min_P d(P, P^{(k)}) . \tag{10}$$

Z definicji mediany Kemeny'ego (10) oraz wzoru (9) wynika, że zadanie wyznaczenia mediany Kemeny'ego można sprowadzić do takiego przestawienia wierszy i kolumn macierzy R (tzn. zmiany kolejności obiektów), aby suma elementów r_{ij} nad przekątną główną była jak najmniejsza. Należy podkreślić, że istnieje wiele sposobów upraszczania zadania wyznaczania mediany Kemeny'ego. Niektóre z nich zostaną omówione w niniejszej pracy.

Definicja 5 [3].

Macierz strat jest przechodnia, jeżeli z warunku

$r_{i_j i_{j+1}} \leq r_{i_{j+1} i_j}$ oraz $r_{i_{j+1} i_{j+s}} \leq r_{i_{j+s} i_{j+1}}$ wynika, że $r_{i_j i_{j+s}} \leq r_{i_{j+s} i_j}$, (11)

$i, = 1,...n$, $j = 1,...n\text{-}1$, $s = 2,\ldots, (n\text{-}j)$.

Warto odnotować, że warunek przechodniości nie wymaga aby s=2. W przypadku ogólnym dla $K \geq 3$ macierz strat odpowiadająca dowolnym uporządkowaniom $P^{(k)} = \{P^1,\ldots,P^K\}$ nie musi być przechodnia.

Definicja 6 [3].

Przyjmujemy, że zwycięzcą w sensie Condorceta jest obiekt, który zdaniem większości ekspertów ($K_W \geq K/2$) poprzedza pozostałe obiekty[2)].

Definicja 7 [3].

Zbiór uporządkowań ma własność Condorceta, jeżeli dla każdego podzbioru obiektów istnieje zwycięzca w sensie Condorceta.

Obiekty będące zwycięzcami w sensie Condorceta są wyznaczane następująco. Najpierw określa się zwycięzcę Condorceta dla całego zbioru obiektów $\{O_i,\}$ i=1, ..., n. Następnie obiekt ten jest usuwany ze zbioru obiektów i wyznaczany jest – o ile istnieje – zwycięzca Condorceta dla (n-1) elementowego podzbioru obiektów. Tak wyznaczony obiekt jest usuwany ze zbioru obiektów i postępowanie jest powtarzane dla kolejnych podzbiorów.

---

2) Niektórzy autorzy przyjmują, że warunek ten ma postać ($K_W > K/2$)

Twierdzenie 1 [3].

Macierz strat odpowiadająca uporządkowaniom $P^{(k)} = \{P^1,\ldots,P^K\}$ jest przechodnia wtedy i tylko wtedy gdy zbiór tych uporządkowań ma własność Condorceta.

Twierdzenie 2 [3].

Jeżeli dany zbiór uporządkowań $P^{(k)} = \{P^1,\ldots,P^K\}$ ma własność Condorceta, to medianę Kemeny'ego stanowi uporządkowanie obiektów $\tilde{P} = \{O_{i_1}, O_{i_2}, \ldots, O_{i_{n-1}}, O_{i_n}\}$ będących zwycięzcami Condorceta dla kolejnych podzbiorów obiektów. Dla tego uporządkowania odległość $d(\tilde{P}, P^{(k)})$ jest kresem dolnym odległości od zbioru uporządkowań $P^{(k)}$. Oznacza się ją literą H.

Twierdzenie 3 [3].

Jeżeli w medianie nie występują obiekty równoważne, wówczas

$$d(P, P^{(k)}) = \sum_{(i,j)\in I_P^{(1)}} r_{ij}\,, \tag{12}$$

gdzie $I_P^{(1)}$ - zbiór indeksów (i,j), dla których w uporządkowaniu P mamy $O_i \succ O_j$, tzn. $a_{ij} = 1$.

Ta zależność może być zastosowana do wyznaczenia kresu dolnego H odległości (9). Litvak wykazał, że spełnione jest następujące twierdzenie.

Twierdzenie 4 [3].

$$H = \sum_{i=1}^{n-1} \sum_{j>i}^{n} min(r_{ij}, r_{ji}) \tag{13}$$

2. Zastosowanie macierzy rozkładu głosów ekspertów

W dotychczasowych rozważaniach posługiwano się przede wszystkim macierzą porównań parami. Wprowadzając do rozważań macierz rozkładu głosów ekspertów można uzyskać istotne wyniki

Założymy, że dany jest zbiór uporządkowań n obiektów podanych przez K ekspertów $P^{(k)} = \{P^1, \ldots, P^K\}$. Założymy również, że w opiniach ekspertów mogą występować obiekty równoważne.

Wprowadzimy oznaczenia:

l_{ij} - liczba ekspertów, których zdaniem $O_i \succ O_j$

l_{ji} - liczba ekspertów, których zdaniem $O_j \succ O_i$ (14)

m_{ij} - liczba ekspertów, których zdaniem $O_i \approx O_j$, to znaczy, że obiekty O_i oraz O_j są równoważne w sensie przyjętego kryterium (zbioru kryteriów).

Zatem spełniona jest równość $l_{ij} + l_{ji} + m_{ij} = K$. (15)

Z definicji macierzy strat $R = [r_{ij}]$ (6), (8) wynika, że

$$r_{ij} = \sum_k \left|a_{ij}^k - 1\right|_{O_i \succ^k O_j} + \sum_k \left|a_{ij}^k - 1\right|_{O_i \approx^k O_j} + \sum_k \left|a_{ij}^k - 1\right|_{O_i \prec^k O_j} \quad (16)$$

skąd

$$r_{ij} = 0 \cdot l_{ij} + 1 \cdot m_{ij} + 2 l_{ji} \quad (17)$$

Jeżeli w uporządkowaniach występują obiekty równoważne zazwyczaj przyjmuje się – aby był spełniony wzór (15) – że każdy z obiektów równoważnych otrzymuje ½ głosu eksperta.

Niech $\tilde{l}_{ij} = l_{ij} + 0.5m_{ij}$. (18)

Wówczas współczynniki macierzy strat można zapisać w postaci

$r_{ij} = 2(l_{ji} + 0{,}5 \cdot m_{ij}) = 2\tilde{l}_{ji}$. (19)

Symbole $\tilde{l}_{ij}$ oraz l_{ij} będą używane zamiennie, ponieważ sposób wyznaczania współczynników macierzy rozkładu głosów ekspertów zazwyczaj wynika z kontekstu. Mamy więc zależność wiążącą macierze strat i rozkładu głosów ekspertów

$R = 2L^T$. (20)

Zależność ta umożliwia sformułowanie podanych przez Litvaka [3] twierdzeń dotyczących własności macierzy strat oraz wyznaczania kresu dolnego odległości za pomocą macierzy rozkładu głosów ekspertów.

Wzór (13) może być zapisany w postaci $H = 2\sum_{i=1}^{n-1}\sum_{j>i}^{n} min(l_{ij}, l_{ji})$ (21)

Odejmując od wyrażenia (17) równość (15) otrzymujemy

$r_{ij} - K = 2l_{ji} + 1 \cdot m_{ij} - l_{ij} - l_{ji} - m_{ij} = l_{ji} - l_{ij} = -\Delta l_{ij}$ (22)

Tak więc $r_{ij} = K - \Delta l_{ij}$, (23)

Ze wzoru (23) wynika także, że $|r_{ij} - r_{ji}| = 2|\Delta l_{ij}|$. (24)

Jeżeli $\Delta l_{ij} \neq 0$, to $|r_{ij} - r_{ji}|_{min} = 2$, bowiem $|\Delta l_{ij}|_{min} = 1$. A zatem minimalna zmiana odległości między uporządkowaniami wynosi 2.

3. Wyznaczanie mediany Kemeny'ego

W pracy Litvaka [3] jest podane twierdzenie mające duże znaczenie praktyczne. W dalszych rozważaniach zostanie pokazane, jak zastosowanie tego twierdzenia pozwala istotnie uprościć proces wyznaczania mediany Kemeny'ego. Różne wersje tego twierdzenia były podawane w kilku pracach, np. Saari, Merlin [4], Truchon [5]. Jednakże ich autorzy nie odwoływali się do oryginalnej pracy Litvaka opublikowanej znacznie wcześniej, bo w 1982 r.

Twierdzenie 5 [3].

Założymy, że jest dany zbiór uporządkowań $P^{(k)} = \{P^1, \dots, P^K\}$, dla którego wyznaczono macierz strat R = $[r_{ij}]$, $i, j = 1, \dots, n$. Jeżeli uporządkowanie $\widetilde{P}$ stanowi medianę tego zbioru uporządkowań, to jest spełniony warunek $r_{i_\nu i_{\nu+1}} \le r_{i_{\nu+1} i_\nu}$, $\nu = 1, \dots, n-1$. (25)

Jeżeli dla danego uporządkowania P jest spełniony warunek (25) a macierz strat R jest przechodnia, to uporządkowanie P stanowi medianę zbioru uporządkowań $P^{(k)}$.

Aby zapisać warunek (25) używając elementów macierzy rozkładu głosów ekspertów, przywołamy wzór (23)

$$r_{i_\nu i_{\nu+1}} = K - (l_{i_\nu i_{\nu+1}} - l_{i_{\nu+1} i_\nu}) \text{ oraz } r_{i_{\nu+1} i_\nu} = K - (l_{i_{\nu+1} i_\nu} - l_{i_\nu i_{\nu+1}}). \tag{26}$$

Warunek (25) można zatem zapisać jak następuje

$$K - (l_{i_\nu i_{\nu+1}} - l_{i_{\nu+1} i_\nu}) \le K - (l_{i_{\nu+1} i_\nu} - l_{i_\nu i_{\nu+1}}), \tag{27}$$

skąd $0 \le 2(l_{i_\nu i_{\nu+1}} - l_{i_{\nu+1} i_\nu})$ (28)

i ostatecznie $0 \le (l_{i_\nu i_{\nu+1}} - l_{i_{\nu+1} i_\nu})$ (29)

Warunek (25) jest warunkiem koniecznym, aby uporządkowanie $\widetilde{P}$ było medianą.

A zatem (29) jest również warunkiem koniecznym, który musi spełniać mediana. Konsekwencją tego warunku jest wymaganie, aby relacje większości między parami obiektów – wynikające z macierzy rozkładu głosów ekspertów – były zachowane w medianie. A zatem na podstawie wzoru (29) można sformułować następujący wniosek.

Uporządkowania, w których nie są spełnione relacje większości wynika-jące z macierzy rozkładu głosów ekspertów, nie mogą być medianą.

Podejście, którego podstawę stanowi to stwierdzenie, zastosował M. Truchon [5] do uproszczenia procesu wyznaczania mediany Kemeny'ego.

Kolejność postępowania jest następująca:

1° Na podstawie macierzy rozkładu głosów ekspertów wyznacza się pary obiektów spełniające relację większości.

2° Dla par obiektów wyznaczonych w punkcie 1° określa się relację przeciwną, tzn., jeżeli między parą obiektów O_i oraz O_j zachodzi relacje większości $O_i \succ O_j$, to przyjmujemy, że $O_j \succ O_i$.

3° Ustalamy wszystkie możliwe uporządkowania, które mogą stanowić medianę. W przypadku n obiektów i przy założeniu, że w medianie nie występują obiekty równoważne, należy rozpatrzyć $n!$ uporządkowań.

4° Ze zbioru uporządkowań wyznaczonych w punkcie 3° usuwamy te, w których między sąsiednimi obiektami występują relacje ustalone w punkcie 2°.

5° Dla uporządkowań, które nie zostały usunięte w kroku 4° sprawdzamy, dla którego uporządkowania (dla których uporządkowań) odległość (9) jest najmniejsza. Uporządkowania, które spełniają ten warunek stanowią medianę Kemeny'ego.

Sprawdzenia, czy jest spełniony warunek z pkt 5° można dokonać na kilka sposobów, w szczególności mając wyznaczoną macierz rozkładu głosów ekspertów L można dla każdego uporządkowania P obliczyć odległość $d(P,P^{(k)})$ zgodnie z wzorem

$$d(P,P^{(k)})=\sum_{i=1}^{n-1}\sum_{j>i}^{n}r_{ij}=\sum_{i=1}^{n-1}\sum_{j>i}^{n}\left[K-(l_{ij}-l_{ji})\right]=\frac{n(n-1)}{2}K-\sum_{i=1}^{n-1}\sum_{j>i}^{n}\Delta l_{ij}\,. \tag{30}$$

W przedstawionej procedurze – służącej do uproszczenia procesu wyznaczania mediany Kemeny'ego – zakłada się eliminację tych uporządkowań, w których sąsiadujące obiekty nie spełniają zasady większości wynikającej z macierzy rozkładu głosów ekspertów.

Jednakże może zdarzyć się sytuacja, że w uporządkowaniach pozostałych po tej eliminacji mogą występować pary obiektów nie zajmujących sąsiednich pozycji, które również nie spełniają relacji większości wynikających z macierzy rozkładu głosów ekspertów. Dla tych par obiektów $\Delta l_{ij}<0$.

A zatem zgodnie z wzorem (30) odległość $d(P,P^{(k)})$ będzie tym mniejsza im suma $\sum_{i=1}^{n-1}\sum_{j>i}^{n}\Delta l_{ij}$ będzie mieć większą wartość. Ze względu na to, że dla rozpatrywanych par obiektów $\Delta l_{ij}<0$, odpowiadające im składniki sumy będą powiększać wartość odległości. Można zatem przyjąć, że uporządkowanie, w którym suma $|\Delta l_{ij}|$ wyznaczona dla par obiektów nie spełniających relacji większości wynikających z macierzy rozkładu głosów ekspertów jest najmniejsza, stanowi medianę Kemeny'ego. Tak właśnie jest formułowana zasada wyznaczania mediany Kemeny'ego w pracach Klamlera [2] oraz Saariego i Merlina [4].

4. Przykłady

Aby zilustrować przedstawiony w pracy algorytm wyznaczania mediany Kemeny'ego, którego podstawę stanowi Twierdzenie 5, podamy jego zastosowanie do konkretnego przykładu.

Przykład 1.

Dany jest zestaw pięciu uporządkowań zbioru czterech obiektów.

$$\begin{aligned} &P^1: O_1, O_2, O_4, O_3\\ &P^2: O_4, O_2, O_1, O_3\\ &P^3: O_2, O_3, O_4, O_1\\ &P^4: O_3, O_4, O_1, O_2\\ &P^5: O_1, O_2, O_3, O_4 \end{aligned} \tag{31}$$

Macierz rozkładu głosów ekspertów

$L=$ (32)

	O_1	O_2	O_3	O_4
O_1	-	3	3	2
O_2	2	-	4	3
O_3	2	1	-	3
O_4	3	2	2	-

Relacja większości jest spełniona jeżeli $l_{ij} \geq 2½$.

Pary obiektów, które spełniają relację większości są następujące:

$O_1 \succ O_2$; $O_1 \succ O_3$; $O_2 \succ O_3$; $O_2 \succ O_4$; $O_3 \succ O_4$; $O_4 \succ O_1$ (33)

Przyjmiemy, że mediany dla zbioru uporządkowań będziemy szukać wśród uporządkowań nie zawierających obiektów równoważnych.

Wiadomo, że dla 4 obiektów takich uporządkowań jest 4!=24. Są one, jak następuje (dla uproszczenia zapisu pominięto symbol $\succ$).

1° O_1 O_2 O_3 O_4	9° O_2 O_3 O_1 O_4	17° O_3 O_4 O_1 O_2	
2° O_1 O_2 O_4 O_3	10° O_2 O_3 O_4 O_1	18° O_3 O_4 O_2 O_1	
3° O_1 O_3 O_2 O_4	11° O_2 O_4 O_1 O_3	19° O_4 O_1 O_2 O_3	
4° O_1 O_3 O_4 O_2	12° O_2 O_4 O_3 O_1	20° O_4 O_1 O_3 O_2	
5° O_1 O_4 O_2 O_3	13° O_3 O_1 O_2 O_4	21° O_4 O_2 O_1 O_3	(34)
6° O_1 O_4 O_3 O_2	14° O_3 O_1 O_4 O_2	22° O_4 O_2 O_3 O_1	
7° O_2 O_1 O_3 O_4	15° O_3 O_2 O_1 O_4	23° O_4 O_3 O_1 O_2	
8° O_2 O_1 O_4 O_3	16° O_3 O_2 O_4 O_1	24° O_4 O_3 O_2 O_1	

Medianą nie mogą być uporządkowania, w których relacje między sąsiednimi obiektami są przeciwne do podanych w (33). Należy zatem wykluczyć uporządkowania, w których

a) $O_1 \succ O_4$; b) $O_2 \succ O_1$; c) $O_3 \succ O_1$;

d) $O_3 \succ O_2$; e) $O_4 \succ O_2$; f) $O_4 \succ O_3$; (35)

Z analizy 24 uporządkowań wynika, że

- warunek a) wyklucza uporządkowania 5°, 6°, 8°, 9°, 14°, 15°
- warunek b) wyklucza uporządkowania 7°, 8°, 15°, 18°, 21°, 24°
- warunek c) wyklucza uporządkowania 9°, 12°, 13°, 14°, 22°, 23
- warunek d) wyklucza uporządkowania 3°, 6°, 15°, 16°, 20°, 24° (36)
- warunek e) wyklucza uporządkowania 4°, 5°, 14°, 18°, 21°, 22°
- warunek f) wyklucza uporządkowania 2°, 6°, 8°, 12°, 23°, 24°

Warunek (33) spełniają jedynie uporządkowania

1° O_1, O_2, O_3, O_4 10° O_2, O_3, O_4, O_1 11° O_2, O_4, O_1, O_3

17° O_3, O_4, O_1, O_2 19° O_4, O_1, O_2, O_3 (37)

Aby wyznaczyć medianę wystarczy sprawdzić, dla którego z tych uporządkowań odległość $d(P, P^{(k)})$ osiąga wartość minimalną lub wyrażenie $\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\Delta l_{ij}$ osiąga wartość maksymalną.

Macierze $[\Delta l_{ij}]$ dla rozpatrywanych uporządkowań mają postać:

	O_1	O_2	O_3	O_4	$\sum_{j>i}\Delta$
O_1	-	1	1	-1	1
O_2		-	3	1	4
O_3			-	1	1
O_4				-	$\boxed{6}$

	O_2	O_3	O_4	O_1	$\sum_{j>i}$
O_2	-	3	1	-1	3
O_3		-	1	-1	0
O_4			-	1	1
O_1				-	$\boxed{4}$

	O_2	O_4	O_1	O_3	$\sum_{j>i}\Delta$
O_2	-	1	-1	3	3
O_4		-	1	-1	0
O_1			-	1	1
O_3				-	$\boxed{4}$

	O_3	O_4	O_1	O_2	$\sum_{j>i}$
O_3	-	1	-1	-3	-3
O_4		-	1	-1	0
O_1			-	1	1
O_2				-	$\boxed{-2}$

	O_4	O_1	O_2	O_3	$\sum_{j>i}\Delta$
O_4	-	1	-1	-1	-1
O_1		-	1	1	2
O_2			-	3	3
O_3				-	4

(38)

Z porównania macierzy $[\Delta l_{ij}]$ wynika, że medianę stanowi uporządkowanie O_1, O_2, O_3, O_4.

Jego odległość od zbioru uporządkowań wyznaczona za pomocą wzoru (30) wynosi $\frac{4(4-1)}{2}5-\sum_i\sum_j \Delta l_{ij}=30-6=24$. Kres dolny macierzy strat określonej dla uporządkowań (31) $H = 22$. Nie istnieje uporządkowanie dla którego $d(P,P^{(k)})=H$. A zatem uporządkowanie O_1, O_2, O_3, O_4 jest medianą, ponieważ odpowiadająca mu odległość od kresu dolnego jest najmniejsza (i równa 2).

Przykład 2.

Załóżmy, że jest dana macierz rozkładu głosów ekspertów wyznaczona na podstawie uporządkowań 5 obiektów, podanych przez 7 ekspertów.

		O_1	O_2	O_3	O_4	O_5
	O_1	0	2.5	3	4	5
$L=$	O_2	4.5	0	3	5	5
	O_3	4	4	0	4.5	5
	O_4	3	2	2.5	0	4.5
	O_5	2	2	2	2.5	0

$$K_W \geq 7/2 = 3.5 \tag{39}$$

Z macierzy L wynika, że obiekt O_3 jest zwycięzcą w sensie Condorceta.

Po wykreśleniu go z macierzy (39) otrzymujemy

	O_1	O_2	O_4	O_5
O_1	0	2.5	4	5
O_2	4.5	0	5	5
O_4	3	2	0	4.5
O_5	2	2	2.5	0

(40)

	O_1	O_4	O_5
O_1	0	4	5
O_4	3	0	4.5
O_5	2	2.5	0

(41)

Kolejnym zwycięzcą w sensie Condorceta jest obiekt O_2.

Po wykreśleniu go z macierzy (40) otrzymujemy macierz (41).

Następnymi zwycięzcami w sensie Condorceta są obiekty O_1 oraz O_4.

Uporządkowanie wynikowe ma postać O_3, O_2, O_1, O_4, O_5. (42)

Macierz $[\Delta l_{ij}]$ uporządkowana w podanej kolejności obiektów jest następująca

	O_3	O_2	O_1	O_4	O_5	$\sum_{j>i}\Delta l_{ij}$
O_3	-	1	1	2	3	7
O_2		-	2	3	3	8
O_1			-	1	3	4
O_4				-	2	2
O_5					-	$\boxed{21}$

(43)

Wykorzystując wzór (30) otrzymujemy, że dla uporządkowania (42) odległość $d(P,P^{(k)}) = 7\frac{5(5-1)}{2} - 21 = 49$.

Kres dolny odległości $d(P,P^{(k)})$ obliczony na podstawie wzoru (21) wynosi 49. A zatem uporządkowanie (42) jest medianą.

Warto podkreślić, że w niektórych przypadkach proces wyznaczania mediany Kemeny'ego można istotnie uprościć, nawet wtedy gdy macierz strat nie jest przechodnia. Sposób postępowania przedstawiono na poniższym przykładzie.

Przykład 3.

Załóżmy, że jest dana macierz rozkładu głosów ekspertów wyznaczona na podstawie uporządkowań 6 obiektów, podanych przez 11 ekspertów.

	O_1	O_2	O_3	O_4	O_5	O_6
O_1	-	7	7	6	6	7
O_2	4	-	7	4	9	6
O_3	4	4	-	4	5	5
O_4	5	7	7	-	7	5
O_5	5	2	6	4	-	5
O_6	4	5	6	6	6	-

(44)

Liczba ekspertów wynosi 11, czyli większość $K_W \geq 11/2 = 5.5$. Pary obiektów (O_i, O_j) spełniające relację większości zacieniowano. Z macierzy (44) wynika, że obiekt O_1 jest zwycięzcą w sensie Condorceta. Po wykreśleniu go z macierzy (44) otrzymujemy

	O_2	O_3	O_4	O_5	O_6
O_2	-	7	4	9	6
O_3	4	-	4	5	5
O_4	7	7	-	7	5
O_5	2	6	4	-	5
O_6	5	6	6	6	-

(45)

W macierzy tej nie występuje zwycięzca w sensie Condorceta, ale przyjmiemy, że obiekt który jest bliski zwycięzcy Condorceta zostaje uznany za tego zwycięzcę. Z macierzy (45) wynika, że może to być obiekt O_2, O_4 lub O_6. Po wykreśleniu obiektu O_4 otrzymujemy)

	O_2	O_3	O_5	O_6
O_2	-	7	9	6
O_3	4	-	5	5
O_5	2	6	-	5
O_6	5	6	6	-

(46)

W macierzy (46) zwycięzcą w sensie Condorceta jest obiekt O_2. Po wykreśleniu go z (46) otrzymujemy

	O_3	O_5	O_6
O_3	-	5	5
O_5	6	-	5
O_6	6	6	-

(47)

W macierzy (47) zwycięzcą w sensie Condorceta jest obiekt O_6.

Po wykreśleniu go z (47) otrzymujemy

	O_3	O_5
O_3	-	5
O_5	6	-

(48)

W macierzy (48) zwycięzcą w sensie Condorceta jest obiekt O_5. Uporządkowanie utworzone z kolejnych zwycięzców w sensie Condorceta ma więc postać

$O_1, O_4, O_2, O_6, O_5, O_3$. (49)

Macierz $[\Delta l_{ij}]$ uporządkowana w podanej kolejności obiektów jest następująca

	O_1	O_4	O_2	O_6	O_5	O_3	$\sum_{j>i} L$
O_1	-	1	3	3	1	3	11
O_4	-1	-	3	-1	3	3	8
O_2	-3	-3	-	1	7	3	11
O_6	-3	1	-1	-	1	1	2
O_5	-1	-3	-7	-1	-	1	1
O_3	-3	-3	-3	-1	-1	-	$\boxed{33}$

(50)

Wykorzystując wzór (30) otrzymujemy, że dla uporządkowania (49) odległość

$$d(P, P^{(k)}) = 11\frac{6(6-1)}{2} - 33 = 132 .$$

Kres dolny odległości $d(P,P^{(k)})$ obliczony na podstawie wzoru (21) wynosi 130. Nie istnieje uporządkowanie dla którego $d(P,P^{(k)})=H$. A zatem uporządkowanie (49) jest medianą, ponieważ odpowiadająca mu odległość od kresu dolnego jest najmniejsza (i równa 2).

Usuwając z macierzy (45) obiekt O_6 otrzymujemy

	O_2	O_3	O_4	O_5
O_2	-	7	4	9
O_3	4	-	4	5
O_4	7	7	-	7
O_5	2	6	4	-

(51)

W macierzy (51) zwycięzcą w sensie Condorceta jest obiekt O_4. Po wykreśleniu go z (51) otrzymujemy

	O_2	O_3	O_5
O_2	-	7	9
O_3	4	-	5
O_5	2	6	-

(52)

Kolejnymi zwycięzcami w sensie Condorceta są obiekty O_2 i O_5. Uporządkowanie utworzone z kolejnych zwycięzców w sensie Condorceta ma więc postać

$$O_1, O_6, O_4, O_2, O_5, O_3. \tag{53}$$

Dla uporządkowania (53) odległość $d(P,P^{(k)})$ wyznaczona za pomocą wzoru (30) wynosi 132 a więc uporządkowanie to jest również medianą.

Usuwając z macierzy (52) obiekt O_2 i następnie obiekty O_6, O_4 i O_5 będące kolejnymi zwycięzcami w sensie Condorceta otrzymujemy uporządkowanie O_1, O_2, O_6, O_4, O_5, O_3. (54)

Dla uporządkowania (54) odległość $d(P,P^{(k)})$ wyznaczona za pomocą wzoru (30) wynosi 136 a więc uporządkowanie to nie jest medianą.

Przykład 4.

Mamy dane uporządkowania 7 obiektów podane przez 11 ekspertów. Są one, jak następuje:

$P^1 : (O_2,O_6),(O_5,O_7),O_3,O_4,O_1$

$P^2 : O_6,(O_1,O_3,O_4),O_5,O_7,O_2$

$P^3 : (O_5,O_6),O_2,(O_1,O_4),(O_3,O_7)$

$P^4 : O_5,(O_1,O_3),(O_2,O_4),O_6,O_7$

$P^5 : O_7,O_4,(O_3,O_6),(O_1,O_2),O_5$

$P^6 : O_4,(O_5,O_6,O_7),(O_1,O_2),O_3$ (55)

$P^7 : O_4,O_5,(O_3,O_7),(O_1,O_2,O_6)$

$P^8 : O_2,O_5,O_7,(O_4,O_6),(O_1,O_3)$

$P^9 : O_6,(O_2,O_5,O_7),(O_1,O_3,O_4)$

$P^{10} : O_1,O_2,O_4,(O_3,O_6),O_5,O_7$

$P^{11} : (O_1,O_2),O_3,(O_4,O_5),(O_6,O_7)$

Liczba głosów wymagana przez zasadę większości $K_W \geq \frac{11}{2} = 5.5$.

Macierz rozkładu głosów z uwzględnieniem połówek głosów dla obiektów równoważnych jest następująca

	O_1	O_2	O_3	O_4	O_5	O_6	O_7
O_1	0	5	6	4.5	4	3.5	5
O_2	6	0	7	7	5.5	5	6.5
O_3	5	4	0	4	4	4	5
O_4	6.5	4	7	0	5.5	6	7
O_5	7	5.5	7	5.5	0	5	8.5
O_6	7.5	6	7	5	6	0	7
O_7	6	4.5	6	4	2.5	4	0

(56)

Pary obiektów, dla których zachodzi relacja większości zacieniowano.

Przyjmiemy, że mediany dla zbioru uporządkowań będziemy szukać wśród uporządkowań nie zawierających obiektów równoważnych.

Wiadomo, że dla 7 obiektów takich uporządkowań jest 7!=5040. Medianą nie mogą być uporządkowania, w których relacje między sąsiednimi obiektami są przeciwne do zacieniowanych. Stosując procedurę wykluczania uporządkowań opisaną w przykładzie 1 otrzymano 9 uporządkowań.

$$\begin{array}{llll|lll|}
O_2, & O_5, & O_4, & O_6, & O_7, & O_1, & O_3 \\
O_2, & O_4, & O_6, & O_5, & O_7, & O_1, & O_3 \\
O_4, & O_6, & O_2, & O_5, & O_7, & O_1, & O_3 \\
O_4, & O_6, & O_5, & O_2, & O_7, & O_1, & O_3 \\
O_5, & O_2, & O_4, & O_6, & O_7, & O_1, & O_3 \\
O_5, & O_4, & O_6, & O_2, & O_7, & O_1, & O_3 \\
O_6, & O_2, & O_4, & O_5, & O_7, & O_1, & O_3 \\
O_6, & O_2, & O_5, & O_4, & O_7, & O_1, & O_3 \\
O_6, & O_5, & O_2, & O_4, & O_7, & O_1, & O_3
\end{array} \tag{57}$$

Dla czterech z nich

$$
\begin{array}{llll|lll|}
\cline{5-7}
O_2, & O_4, & O_6, & O_5, & O_7, & O_1, & O_3 \\
O_6, & O_2, & O_5, & O_4, & O_7, & O_1, & O_3 \\
O_6, & O_5, & O_2, & O_4, & O_7, & O_1, & O_3 \\
O_6, & O_2, & O_4, & O_5, & O_7, & O_1, & O_3 \\
\cline{5-7}
\end{array}
\tag{58}
$$

odległość $d(P, P^{(k)}) = 11\frac{7(7-1)}{2} - 43 = 188$.

Kres dolny odległości $d(P, P^{(k)})$ obliczony na podstawie wzoru (30) wynosi 186 i nie istnieje uporządkowanie dla którego $d(P, P^{(k)}) = H$. A zatem uporządkowania (58) są medianą, ponieważ odpowiadająca im odległość od kresu dolnego jest najmniejsza (i równa 2).

Uwagi końcowe

Zależność (20) wiążąca macierz strat R z macierzą rozkładu głosów ekspertów L umożliwia bezpośrednie zastosowanie tej ostatniej do wyznaczania mediany Kemeny'ego. Warunek konieczny (29), aby – wynikające z macierzy L – relacje większości między obiektami były zachowane w medianie, umożliwia eliminację ze zbioru rozważanych $n!$ uporządkowań tych uporządkowań, w których te relacje nie są spełnione. Podejście to w istotny sposób ogranicza liczność zbioru uporządkowań przeszukiwanych przy stosowaniu przeglądu zupełnego i może być wykorzystane do efektywnego wyznaczania mediany Kemeny'ego.

Ponadto, w niektórych przypadkach, nawet gdy zbiór danych uporządkowań nie ma własności Condorceta, proces wyznaczania mediany Kemeny'ego można znacząco uprościć przyjmując obiekt bliski zwycięzcy Condorceta za tego zwycięzcę w celu utworzenia ciągu kolejnych zwycięzców Condorceta.

Zastosowanie macierzy rozkładu głosów ekspertów L jest istotne również z uwagi na fakt, że niekiedy macierz ta stanowi jedyne źródło informacji o preferencjach ekspertów (macierze porównań parami mogą nie być dostępne).

Literatura

[1] Kemeny J.G., Snell L.J., *Preference Ranking: An Axiomatic Approach.* In J.G. Kemeny and L.J. Snell, Mathematical Models in the Social Sciences, New York, Ginn, 1962.

[2] Klamler C., 2004, *The Dodgson ranking and its relation to Kemeny's method and Slater's rule*. Social Choice and Welfare **23**, 91–102.

[3] Litvak B.G., *Ekspertnaja informacija. Mietody połuczienija i analiza,* Radio i Swjaz, Moskwa,1982.

[4] Saari, D.G., Merlin V.R., 2000, *A geometric examination of Kemeny's rule*. Social Choice and Welfare **17**, 403–438.

[5] Truchon M., 2004, *Aggregation of Rankings in Figure Skating,* Cahier de recherche 04-14, CIPANO, CIRPEE and Departement d'economique, Universite Laval, Quebec, Canada.