Modelovanje preferencija – metoda Promethee

Vrednosti u matrici odlučivanja često predstavljaju direktno izmerene vrednosti i kao takve ne odražavaju koliko je DO određena vrednost *korisna* ili *povoljna*[1]. Stoga, te izmerene vrednosti (npr. brzina automobila ili cena kafe) se mogu mapirati u vrednosti koje predstavljaju *korist* te vrednosti. Jedan način za modelovanje korisnosti je preko Teorije korisnosti (pogledati predavanja). Međutim, određivanje korisnosti može biti komplikovano, te DO može izraziti svoje mišljenje preko relacije da je jedna alternativa "bolja od" druge alternative, za dati kriterijum; tj. DO "preferira" jednu alternativu u odnosu na drugu alternativu, za dati kriterijum. Modelovanje preferencija se vrši tako što se porede sve moguće kombinacije alternative, za svaki kriterijum, gde svaki od kriterijuma ima definisanu funkciju preferencije. Ukoliko *p(A, B)* predstavlja vrednost preferencije alternative A, u odnosu na alternativu B, onda kažemo:

- Ako je p(A, B) = 0, onda ne preferiramo alternativu A, u odnosu na alternativu B (treba voditi računa da ova vrednost ne znači da preferiramo alternativu B, u odnosu na alternativu A).
- Ako je p(A, B) ≈ 0, onda iskazujemo slabu preferencu prema alternativi A, u odnosu na alternativu B.
- Ako je p(A, B) ≈ 1, onda jako preferiramo alternativu A, u odnosu na alternativu B.
- Ako je p(A, B) = 1, onda u potpunosti preferiramo alternativu A, u odnosu na alternativu B.

Treba imati u vidu da je p(A, B) uvek pozitivan broj. Način na koji se dolazi do funkcija preferencija je putem anketa, upitnika ili ispitivanjem DO. Na ovaj način se uvodi subjektivnost DO.

Primer:

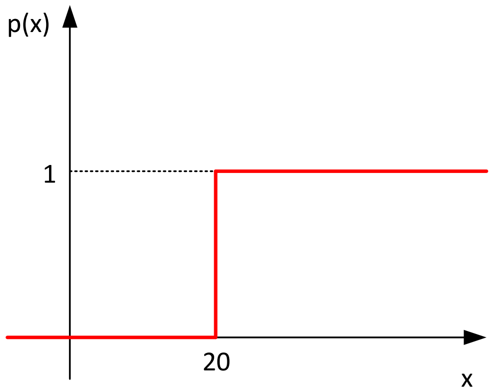
Pre odlaska na fakultet želimo da popijemo kafu, te smo strukturirali problem na sledeći način:

	Cena (min)	Kvalitet (max)
A	100	3
B	150	4
C	110	5

Ukoliko pogledamo alternative A i C, vidimo da je razlika među njima 10 dinara. Međutim, kako je razlika relativno niska, ne želimo da preferiramo alternativu A u odnosu na alternativu C. Kako bismo iskazali da nam određena vrednost ne predstavlja značajnu razliku, kreiramo funkciju

[1] vrednosti u matrici odlučivanja koje su procenjene donekle predstavljaju korisnost DO, jer predstavljaju mišljenje DO o alternativi, izraženo, na primer, na skali od 1 do 9

preferencije. Recimo da razlika koja nam nije bitna iznosi 20 dinara. U tom slučaju, kreiramo sledeću funkciju:



Vrednost na x osi predstavlja razliku vrednosti prilikom poređenja dve alternative, a vrednost na y osi predstavlja vrednost preferencije. Ukoliko je razlika vrednosti dve alternative manja od ili jednaka 20, onda kažemo da ne preferiramo alternativu A u odnosu na alternativu B. Odnosno, ako je razlika veća od 20, onda kažemo da preferiramo alternativu A u odnosu na alternativu B. Vrednost 20 (ili m u opštem slučaju) predstavlja parametar indiferencije ili ravnodušnosti, koji označava vrednost u razlici koja DO nije bitna, tj. DO je ravnodušan prema toj razlici.

Drugim rečima:

$$p(x) = \begin{Bmatrix} 0, x \leq 20 \\ 1, x > 20 \end{Bmatrix}$$

tj. u opštem slučaju

$$p(x) = \begin{Bmatrix} 0, x \leq m \\ 1, x > m \end{Bmatrix}$$

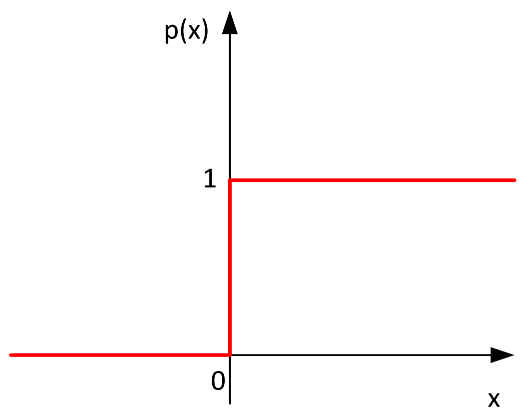
Ova funkcija se naziva funkcijom preferencije. Konkretno, ova funkcija preferencije predstavlja drugi tip preferencije. Tipova preferencije ima neograničeno mnogo jer svaka funkcija koja razliku dva broja preslikava u vrednosti između nula i jedan predstavlja funkciju preferencije. U knjizi je objašnjeno sedam najčešće korišćenih tipova funkcija preferencije.

Imajući u vidu datu funkciju, možemo izraziti preferencije. Treba napomenuti da računanje razlike zavisi od tipa ekstremizacije. Stoga, ukoliko računamo preferenciju alternative A u odnosu na alternativu B, razlike računamo na sledeći način:

	A vs. B
max	A - B
min	-(A - B) = B - A

	x	Cena p(x)
(A, B)	150 – 100 = 50 > 20	**1**
(A, C)	110 – 100 = 10 ≤ 20	0
(B, A)	100 – 150 = -50 ≤ 20	0
(B, C)	110 – 150 = -40 ≤ 20	0
(C, A)	100 – 110 = -10 ≤ 20	0
(C, B)	150 – 110 = 40 > 20	**1**

Za kvalitet usluge kafića, možemo da smatramo da nam je svaka razlika u vrednostima veoma bitna. Ovo je prvi tip preferencije. Drugim rečima, parametar indiferencije je jednak 0.



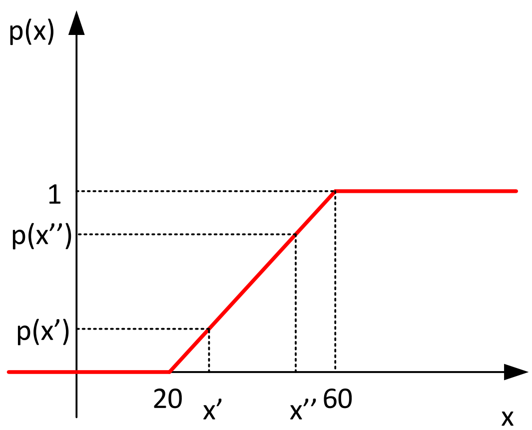
U tom slučaju, vrednost preferencije se dobija na sledeći način.

$$p(x) = \begin{Bmatrix} 0, x \leq 0 \\ 1, x > 0 \end{Bmatrix}$$

	x	Kvalitet p(x)
(A, B)	3 – 4 = -1 ≤ 0	0
(A, C)	3 – 5 = -2 ≤ 0	0
(B, A)	4 – 3 = 1 > 0	**1**
(B, C)	4 – 5 = -1 ≤ 0	0
(C, A)	5 – 3 = 2 > 0	**1**
(C, B)	5 – 4 = 1 > 0	**1**

Međutim, često DO ne može striktno da preferira neku alternativu u odnosu na drugu, tj. ne može da odredi vrednosti koje su tačno jednake 0 ili 1, već želi da predstavi preferenciju kao vrednost koja može biti između 0 ili 1. Na primeru kafića, DO želi da kaže da je indiferentan prema razlici od 20 dinara, ali isto tako, da tek kad razlika iznosi 60 dinara, smatra da u potpunosti preferira prvu alternativu u odnosu na drugu. Ako je razlika između 20 i 60 dinara, onda je vrednost funkcije preferencije između 0 i 1, i to: ako je razlika bliže 20, onda je vrednost funkcije preferencije bliža nuli, a ako je razlika bliža 60, onda je vrednost funkcije preferencije bliža 1.

Funkcija preferencije koja modeluje ovu situaciju predstavlja peti tip (opšti tip) preferencije i definiše se preko parametra indiferencije *m*, koji predstavlja ravnodušnost DO prema razlici, i parametra preferencije *n*, koji predstavlja vrednost razlike posle koje DO u potpunosti preferira jednu alternativu u odnosu na drugu. Za dati primer, funkcija preferencije je definisana na sledeći način.



Vrednost funkcije preferencije se dobija na sledeći način:

$$p(x) = \begin{Bmatrix} 0, x \leq 20 \\ \dfrac{x-20}{60-20}, 20 < x \leq 60 \\ 1, x > 60 \end{Bmatrix}$$

tj. u opštem slučaju

$$p(x) = \begin{Bmatrix} 0, x \leq m \\ \dfrac{x-m}{n-m}, m < x \leq n \\ 1, x > n \end{Bmatrix}$$

Koristeći ovu funkciju preferencije, dobijamo sledeće vrednosti preferencija:

	x	Cena p(x)
(A, B)	150 – 100 = 50	$\frac{50-20}{60-20}=\frac{30}{40}=\frac{3}{4}$
(A, C)	110 – 100 = 10 ≤ 20	0
(B, A)	100 – 150 = -50 ≤ 20	0
(B, C)	110 – 150 = -40 ≤ 20	0
(C, A)	100 – 110 = -10 ≤ 20	0
(C, B)	150 – 110 = 40	$\frac{40-20}{60-20}=\frac{20}{40}=\frac{1}{2}$

Prema redosledu pojavljivanja, do sada smo naveli: drugi, prvi i peti tip preferencije. Postoji i treći tip (linearni tip), odnosno:

$$p(x)=\begin{Bmatrix} 0, x \le m \\ \frac{x}{n}, m < x \le n \\ 1, x > n \end{Bmatrix}$$

Pored navedenih, postoji i četvrti tip preferencije, koji je definisan:

$$p(x)=\begin{Bmatrix} 0, x \le m \\ \frac{1}{2}, m < x \le n \\ 1, x > n \end{Bmatrix}$$

Metoda se rešava preko sledećih koraka:

1. Računanje preferencija po parovima alternativa, za svaki kriterijum,
2. Računanje težinski sabranih preferenci, po parovima alternativa,
3. Računanje pozitivnog i negativnog toka alternativa,
4. Računanje čistog toka alternativa.

Zadatak:

Napravili ste e-prodavnicu i potrebno je da izaberete server na kome ćete okačiti aplikaciju. Nakon istraživanja tržišta i traganja za alternativama došli ste do sledeće matrice odlučivanja.

	Cena (min)	Opterećenost [%] (min)	Pouzdanost (max)
Srbija	50	35	4
Kina	20	100	2
Hrvatska	70	30	2
Amerika	75	60	5
Rumunija	80	70	3

Ponderi	0,4	0,3	0,3

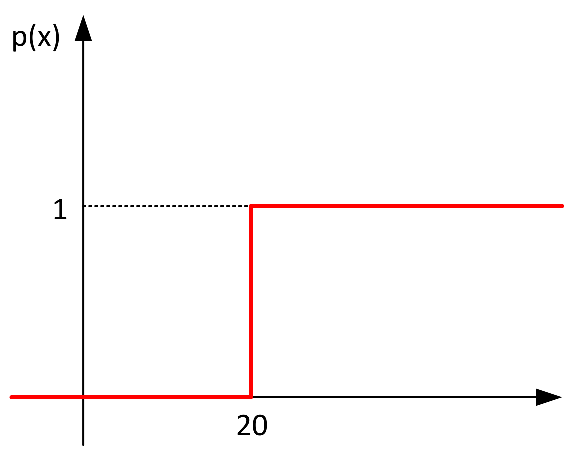
Indiff. (*m*)	20	10	0
Pref. (*n*)	20	40	0

Posmatrajući matricu odlučivanja, možemo da zaključimo da je Rumunija dominirana alternativa, te je možemo isključiti iz matrice odlučivanja.

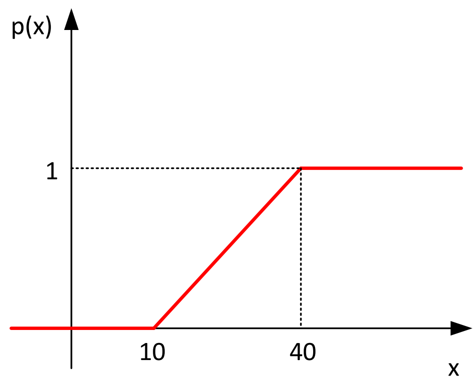
Funkcije preferencije su predstavljene preko parametara *m* i *n*, a važnost kriterijuma je data preko pondera.

1. **Računanje preferencija po parovima alternativa za svaki kriterijum**

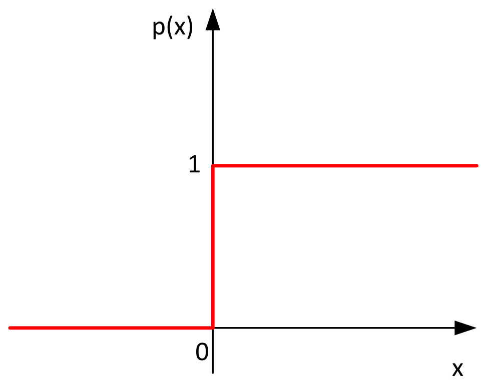
Za kriterijum Cena, funkcija preferencije izgleda:



Za kriterijum Opterećenost, funkcija preferencije izgleda:



Za kriterijum Pouzdanost, funkcija preferencije izgleda:



	Cena (min)	Opt. [%] (min)	Pouz. (max)
Srbija, Kina	0	1	1
Srbija, Hrvatska	0	0	1
Srbija, Amerika	1	$\frac{25-10}{40-10}=\frac{1}{2}$	0
Kina, Srbija	1	0	0
Kina, Hrvatska	1	0	0
Kina, Amerika	1	0	0
Hrvatska, Srbija	0	0	0
Hrvatska, Kina	0	1	0
Hrvatska, Amerika	0	$\frac{30-10}{40-10}=\frac{2}{3}$	0
Amerika, Srbija	0	0	1
Amerika, Kina	0	1	1
Amerika, Hrvatska	0	0	1

2. **Računanje težinski sabranih preferenci po parovima alternativa**

Ovaj korak se sprovodi tako što se za svaki kriterijum saberu proizvodi težina (pondera) i vrednosti funkcije preferencije. Odnosno:

$P_{ik} = \sum_{j=1}^{n} t_j * P_j(a_i, a_j)$

	Preferencija
Srbija, Kina	0,4 * 0 + 0,3 * 1 + 0,3 * 1 = **0,6**
Srbija, Hrvatska	0,4 * 0 + 0,3 * 0 + 0,3 * 1 = **0,3**
Srbija, Amerika	0,4 * 1 + 0,3 * 0,5 + 0,3 * 0 = **0,55**
Kina, Srbija	0,4 * 1 + 0,3 * 0 + 0,3 * 0 = **0,4**
Kina, Hrvatska	0,4 * 1 + 0,3 * 0 + 0,3 * 0 = **0,4**
Kina, Amerika	0,4 * 1 + 0,3 * 0 + 0,3 * 0 = **0,4**
Hrvatska, Srbija	0,4 * 0 + 0,3 * 0 + 0,3 * 0 = **0**
Hrvatska, Kina	0,4 * 0 + 0,3 * 1 + 0,3 * 0 = **0,3**
Hrvatska, Amerika	0,4 * 0 + 0,3 * 2/3 + 0,3 * 0 = **0,2**
Amerika, Srbija	0,4 * 0 + 0,3 * 0 + 0,3 * 1 = **0,3**
Amerika, Kina	0,4 * 0 + 0,3 * 1 + 0,3 * 1 = **0,6**
Amerika, Hrvatska	0,4 * 0 + 0,3 * 0 + 0,3 * 1 = **0,3**

Ovako dobijene vrednosti iskazuju koliko ukupno preferiramo alternativu A, u odnosu na alternativu B. Npr. Kinu, u odnosu na Ameriku, ukupno preferiramo 0,4 (40%).

Dobijene preferencije treba preneti u matricu datu u sledećem koraku.

3. **Računanje pozitivnog i negativnog toka alternativa**

Pozitivan tok preferencije predstavlja prosečnu vrednost reda, i govori koliko, u proseku, preferiramo posmatranu alternativu u odnosu na druge.

Negativan tok preferencije prestavlja prosečnu vrednost kolone, i govori koliko su prosečno druge alternative bile preferirane u odnosu na posmatranu.

Dakle, pozitivan tok alternative Srbija, dobija se kao (0,6 + 0,3 + 0,55)/3. Negativni tok se dobija kao (0,4 + 0 + 0,3)/3.

	Srbija	Kina	Hrvatska	Amerika	T^+
Srbija		0,6	0,3	0,55	0,483
Kina	0,4		0,4	0,4	0,4
Hrvatska	0	0,3		0,2	0,167
Amerika	0,3	0,6	0,3		0,4
T^-	0,233	0,5	0,333	0,383	

4. **Računanje čistog toka alternativa**

Čist tok (ili neto tok) preferencije predstavlja razliku pozitivnog i negativnog toka. Suma čistog toka treba da bude 0. Veća vrednost čistog toga predstavlja bolju vrednost, te zaključujemo da je najbolja alternativa ona koja ima najveću vrednost.U našem primeru, to je alternativa Srbija.

	Srbija	Kina	Hrvatska	Amerika	T^+	T
Srbija		0,6	0,3	0,55	0,483	**0,25**
Kina	0,4		0,4	0,4	0,4	-0,1
Hrvatska	0	0,3		0,2	0,167	-0,167
Amerika	0,3	0,6	0,3		0,4	0,017
T^-	0,233	0,5	0,333	0,383		Suma = 0

Jednokriterijumska analiza preferencija

Promethee metoda može da se koristi u cilju jednokriterijumske analize preferencija, odnosno, kako bi se dobile vrednosti preferencija alternativa za jedan kriterijum. Intuitivno je jasno da dobijene vrednosti moraju biti u opsegu [-1, 1], a da je suma svih vrednosti nula.

Primer:

Potrebno je izabrati automobil u zavisnosti od kriterijuma Cena. DO ne smatra svaku razliku u ceni podjednako važnom, već smatra da razlika do 2000 evra nije bitna. Zatim, što je razlika u ceni veća, to je važnost proporcionalno viša, sve do granice od 5000 evra. Nakon te granice, svaka razlika je u potpunosti bitna.

Dakle, imamo:

Alternative kola	**Cena**
Ekonomska klasa	15000
Sportska	29000
Luksuzna	38000
Turing A	24000
Turing B	25500

Takođe imamo peti, opšti tip preferencije:

Indiff. (*m*)	2000
Pref. (*n*)	5000

Koraci rešavanja su slični. Prva razlika u rešavanju nalazi se u prvom koraku, gde se korak sprovodi samo za kriterijum koji posmatramo. Zatim, kako računamo preferencije samo za jedan kriterijum, preskačemo drugi korak (težinski sabrane preferencije). Dakle, prvo računamo preferencije po parovima alternativa:

	Ekonomska klasa	Sportska	Luksuzna	Turing A	Turing B
Ekonomska klasa		1	1	1	1
Sportska	0		1	0	0
Luksuzna	0	0		0	0
Turing A	0	1	1		0
Turing B	0	0,5	1	0	

Zatim, računamo pozitivan i negativan tok preferencija.

	Ekonomska klasa	Sportska	Luksuzna	Turing A	Turing B	T^+
Ekonomska klasa		1	1	1	1	1
Sportska	0		1	0	0	0,25
Luksuzna	0	0		0	0	0
Turing A	0	1	1		0	0,5
Turing B	0	0,5	1	0		0,375
T^-	0	0,625	1	0,25	0,25	

Odavde računamo čist tok preferencija.

	T^+	T^-	T
Ekonomska klasa	1	0	1
Sportska	0,25	0,625	-0,375
Luksuzna	0	1	-1
Turing A	0,5	0,25	0,25
Turing B	0,375	0,25	0,125

Ovako dobijene vrednosti predstavljaju preferencije alternativa DO za kriterijum Cena. Kao takve, mogu se koristiti u tabeli odlučivanja. Odnosno, jednokriterijumska analiza preferencija predstavlja jedan od načina na koji se Promethee metoda može koristiti za popunjavanje tabele odlučivanja i za kombinovanje Promethee metode sa drugim metodama.

Primer:

Donosilac odluke (DO) želi da nabavi novi mobilni telefon, koji je spreman da plati između 400 i 700 eura. Prilikom izbora alternative, cena je najvažniji kriterijum, ali je DO spreman da plati nešto višu cenu da bi dobio bolje karakteristike. Nakon inicijalne selekcije alternativa, DO formirao je sledeću tabelu odlučivanja.

	Cena (eur) min	Procesor (Broj jezgara) max	Memorija (GB) max	Kamera (mpx) max
A	500	2	32	12
B	700	4	64	16
C	460	2	32	4
Težine	**0,6**	**0,1**	**0,2**	**0,1**

DO je nad datom tabelom odlučivanja sproveo JAT metodu, pri čemu je normalizovao tabelu "IKOR" normom. Normalizovana tabela, kao i vrednosti očekivane koristi svake alternative, date su u tabeli ispod. Možemo primetiti da je, prema odluci/rešenju JAT metode, najbolja "alternativa C" (očekivana korist – 0.6).

Ukoliko detaljnije analiziramo vrednosti alternativa po svakom kriterijumu, možemo primetiti da "alternativa B" ima sve karakteristike bolje u odnosu na druge alternative, ali je drastično skuplja (700 eura u odnosu na 500 i 460). Kako je DO-u cena najvažniji kriterijum (težina 0.6), očekivano je da ova alternativa zauzima poslednje mesto po očekivanoj koristi. Sa druge strane, "alternativa A" i "alternativa C" su podjednako dobre po kriterijumima "Procesor" i "Memorija", dok je alternativa A" drastično bolja u odnosu na "alternativu C" po kriterijumu kamera (12 megapiksela u odnosu na 4 megapiksela). Uprkos boljoj kameri, razlika u ceni od 40 eura je doprinela tome da "alternativa C", prema JAT metodi, ima lošiju očekivanu korist u odnosu na "alternativu A", obzirom da je "Cena" najvažniji kriterijum.

	Cena (eur) **min**	Procesor (Broj jezgara) **max**	Memorija (GB) **max**	Kamera (mpx) **max**	**OK** **max**
A	0,833	0	0	0,667	0,567
B	0	1	1	1	0,4
C	1	0	0	0	0,6
Težine	**0,6**	**0,1**	**0,2**	**0,1**	

Nakon inicijalne analize rešenja, DO je shvatio da je on indiferentan prema razlici u ceni koja je manja od 50 eura, tj. DO bi platio 50 eura zarad izbora kvalitetnije alternative. Prema tome, DO je rešio da modeluje kriterijum "Cena" funkcijom preferencije, pri čemu su parametri indiferencije i preferencije jednaki 50.

Indiff. (*m*)	50
Pref. (*n*)	50

Na osnovu početne tabele odlučivanja, izračunate su preferencije, a na osnovu poređenja alternativa po parovima.

Cena	A	B	C
A	-	1	0
B	0	-	0
C	0	1	-

Na osnovu preferencija iz prethodne tabele, izračunat je pozitivan, negativan i čist tok preferencije (tabela ispod).

Cena	T^+	T^-	T
A	0.5	0	0.5
B	0	1	-1
C	0.5	0	0.5

Kao što je rečeno ranije, čist tok preferencije uzima vrednosti između -1 i 1, pri čemu 1 predstavlja najbolju vrednost, a -1 najlošiju (zato što se čist tok dobija tako što se od pozitivnog toga oduzima negativan tok). Prema tome, čist tok preferencije se može posmatrati kao kriterijum "Cena" koji je tipa maksimizacije i koji se može koristiti u JAT (ili bilo kojoj drugoj metodi), tako što originalni kriterijum (iz početne tabele ovog primera) menjamo čistim tokom preferencije.

Nakon zamene početnog kriterijuma "Cena" čistim tokom preferencije, u kojem je integrisana indiferentnost prema razlikama u ceni koje su manje od 50 eura, tabela odlučivanja izgleda ovako:

	Cena (eur) **min**	Procesor (Broj jezgara) **max**	Memorija (GB) **max**	Kamera (mpx) **max**
A	0,5	2	32	12
B	-1	4	64	16
C	0,5	2	32	4
Težine	**0,6**	**0,1**	**0,2**	**0,1**

Na osnovu tabele, možemo zaključiti da je integracija čistog toka preferencije dovela do željenog ishoda: “alternativa A” i “alternativa B” su podjednako dobre (0.5), dok je “alternativa C” lošija. Ovim postupkom smo u problem odlučivanja integrisali intuiciju da je donosilac odluke indiferentan prema razlici od 50 eura, kada su cene alternativa u rasponu između 460 i 700 eura.

Sa druge strane, pozitivan tok može da uzima negativne vrednosti (u ovom slučaju -1 za “alternativu B”) i to predstavlja problem za računanje i pravilno tumačenje očekivane koristi (otežane sume vrednosti alternativa po kriterijumima). Takođe, kriterijum formiran na osnovu čistog toka preferencije nije uporediv sa drugim kriterijumima zbog različitog reda veličine koje uzima u odnosu na ostale kriterijume. Zbog toga je, pre računanja očekivane koristi, neophodno izvršiti normalizaciju svih kriterijuma koristeći istu normu. Korišćenjem L1 (deljenje zbirom) ili L∞ (deljenje maksimalnom vrednošću) norme, zadržali bismo negativne vrednost iz tabele odlučivanja. Da bismo ovo izbegli, a pri tom zadržali odnos vrednosti između alternativa po kriterijumu cena, možemo koristiti “IKOR” normu. Ova norma nije osetljiva na negativne vrednosti, zbog toga što umesto vrednosti alternativa po nekom kriterijumu, posmatra normalizovano odstojanje vrednosti alternative od najbolje alternative. Kako se normalizacija odstojanja vrši deljenjem razlikom između najbolje i najlošije alternative, normalizovane vrednosti će uvek biti pozitivne i tipa maksimizacije. Tabela ispod prikazuje vrednosti alternativa normalizovane “IKOR” normom, kao I očekivanu korist izračunatu na osnovu tih vrednosti i težina.

	Cena (eur) **min**	Procesor (Broj jezgara) **max**	Memorija (GB) **max**	Kamera (mpx) **max**	**OK** **max**
A	1	0	0	0,667	0,667
B	0	1	1	1	0,4
C	1	0	0	0	0,6
Težine	**0,6**	**0,1**	**0,2**	**0,1**	

U ovom slučaju, “alternativa A” se pokazala kao najbolja. Štaviše, nakon zanemarivanja razlike u ceni od 40 eura, “alternativa C” postaje dominirana u odnosu na “alternativu A”, obzirom da je „alternativa A“ po svim kriterijumima jednaka, a po kriterijumu “Kamera” bolja od „alternative C“.

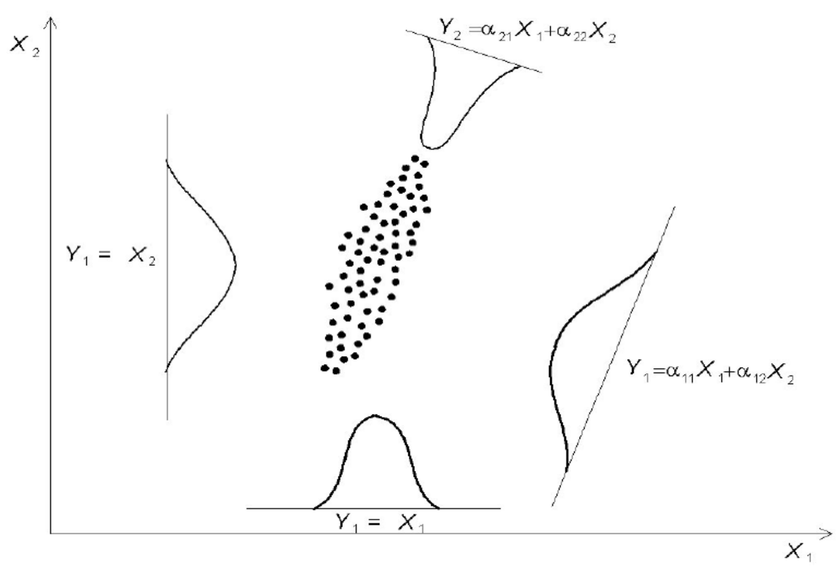
Promethee i analiza glavnih komponenti

Analiza glavnih komponenti može da se koristi da objasni varijansu svih kriterijuma (ispravnije je reči što više varijanse) u tabeli odlučivanja. Kao rezultat, dobijaju se komponente, koje predstavljaju linearnu kombinaciju svih kriterijuma u tabeli odlučivanja. Dobijene komponente su takve da prva komponenta objašnjava najveći deo varijabiliteta. Zatim, druga je nezavisna od prve (korelacija sa prvom komponentom je jednaka 0), a pritom uzima najveći deo preostalog varijabiliteta, itd.

Primer glavnih komponenti je prikazan na slici ispod. Promenjive X1 i X2 su opisane preko dve nove promenjive Y1 i Y2, od kojih svaka predstavlja linearnu kombinaciju X1 i X2 (primerite formulu na slici pored Y1 i Y2). One, kao takve, objašnjavaju raspršenost u podacima (varijansu, tj. varijabilitet) i nekorelisane su jedna sa drugom (primetite da su Y1 i Y2 normalne jedna u poređenju sa drugom, odnosno, da je ugao između njih 90 stepeni). U opštem slučaju, možemo imati više promenjivih X, a ideja je da korišćenjem glavnih komponenti dobijemo manji broj Y koji će objasniti što više podataka (što više varijanse). Korišćenje prve dve glavne komponente je veoma pogodno iz dva razloga:

1) problem se može grafički predstaviti i

2) te komponente objašnjavaju najviše varijanse.



Slika 1. Projekcija komponenti u skupu podataka[2]

Analiza glavnih komponenti je pogodna za kombinovanje sa Promethee metodom jer se lepo uklapa sa matricom preferencija. Naime, analiza glavnih komponenti u kombinaciji sa

[2] Kovačić Z, (1994) Multivarijaciona analiza, Ekonomski fakultet, Beograd

Promethee metodom se naziva GAIA[3] i predstavlja vizualizaciju alternativa i kriterijuma na dvodimenzionu ravan, čime se omogućava da DO vidi koje alternative su „jake“ za koje kriterijume, kao i koliko su alternative međusobno slične. Napominjemo da se analiza glavnih komponenti može koristiti is a drugim metodama, ne samo sa Promethee metodom.

Koraci za primenenu analize glavnih komponenti i Promethee metode su:

1. Računanje preferencija po parovima alternativa, za svaki kriterijum,
2. Konstrukcija matrice čistih tokova (računanje čistog toka za svaki kriterijum ponaosob),
3. Računanje matrice kovarijanse,
4. Računanje sopstvenih vrednosti i sopstvenih vektora kovarijacione matrice (matrice kovarijanse),
5. Projektovanje podataka na prve dve glavne komponente.

Primer:

Imamo tri kriterijuma i tri alternative zadate na sledeći način:

	K1 (min)	K2 (max)	K3 (min)
A1	12	45	65
A2	15	79	63
A3	23	28	35

Ponderi	0,3	0,5	0,2

Indiff. (*m*)	0	30	0
Pref. (*n*)	0	30	45

Prvo računamo, kao i kod izvorne Promethee metode, preferencije po parovima alternativa, za svaki kriterijum, te dobijamo matricu preferencija:

	K1	K2	K3
A1, A2	1	0	0
A1, A3	1	0	0
A2, A1	0	1	2/45
A2, A3	1	1	0
A3, A1	0	0	30/45
A3, A2	0	0	28/45

[3] Više informacija o upotrebi i tumačenju GAIA ravni u softveru, može se pronaći na linku. Potrebno je preuzeti uputstvo za upotrebu.

Zatim, računamo čist tok, za svaki kriterijum ponaosob. Odnosno, prvo računamo pozitivan i negativan tok, pa onda čist tok. Za kriterijum K1, dobijamo:

K1	A1	A2	A3	T^+	T
A1		1	1	1	1
A2	0		1	0,5	0
A3	0	0		0	-1
T^-	0	0,5	1		Suma = 0

Za kriterijum K2, dobijamo:

K2	A1	A2	A3	T^+	T
A1		0	0	0	-0,5
A2	1		1	1	1
A3	0	0		0	-0,5
T^-	0,5	0	0,5		Suma = 0

Na kraju, za kriterijum K3, dobijamo:

K3	A1	A2	A3	T^+	T
A1		0	0	0	-0,356
A2	0,044		0	0,22	-0,289
A3	0,667	0,622		0,645	0,645
T^-	0,356	0,311	0		Suma = 0

Kada smo izračunali čiste tokove, sastavljamo matricu čistih tokova:

	K1	K2	K3
A1	1	-0,5	-0,356
A2	0	1	-0,289
A3	-1	-0,5	0,645

Iz matrice čistih tokova prvo računamo matricu kovarijanse. Formula je: $\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \acute{x})(y_i - \acute{y})$, gde su x i y kolone matrice čistog toka, a i red u matrici. Vrednosti $\acute{x}$ i $\acute{y}$ su prosečne vrednosti kolone. Na našem primeru, dobijamo:

Kovarijansa	K1	K2	K3
K1	1	0	-0,501
K2	0	0,75	-0,217
K3	-0,501	-0,217	0,313

Iz matrice kovarijansi računamo matricu sopstvenih vrednosti. Ovaj korak nam je neophodan kako bismo konstruisali GAIA ravan, te projektovali kriterijume i alternative na nju. Za računanje matrice sopstvenih vrednosti koristite sledeći link. Dobijamo sledeće sopstvene vrednosti i matricu sopstvenih vektora:

λ.	1,284	0	0,779

	v1	v2	v3
K1	0,853	-0,433	0,291
K2	0,197	-0,250	-0,948
K3	-0,484	0,866	0,128

Sopstvene vrednosti predstavljaju intenzitet varijabliteta koju opisuje sopstveni vektor. Prva glavna komponenta ima najveću vrednost sopstvenog vektora (u našem primeru to je λ1, sa vrednošću 1,284), zatim druga glavna komponenta ima drugu najveću vrednost sopstvenog vektora (u našem primeru to je λ3, sa vrednošću 0,779), itd. Varijabilitet koji se objašnjava preko novih glavnih komponenti se računa kao udeo sopstvene vrednosti svake komponente, u ukupnom zbiru sopstvenih vrednosti. Odnosno, dobijamo da nam prva komponenta koja se dobija iz λ1 objašnjava $\frac{1,284}{1,284+0+0,779} = 62,24$, a druga glavna komponenta $\frac{0,779}{1,284+0+0,779} = 37,76$varijabiliteta. Za kreiranje GAIA ravni potrebne su nam prve dve glavne komponente. U ovom primeru, prve dve komponente pokrivaju 100 varijabiliteta (62,24% + 37,76% = 100%), odnosno, ne postoji gubitak informacije što predstavlja idealan slučaj. U opštem slučaju, smatra se da je problem dobro prikazan preko GAIA ravni ukoliko objašnjava 80% varijabiliteta, odnosno da je dovoljno dobro prikazan ukoliko objašnjava 60% varijabiliteta.

Sada podatke projektujemo na glavne komponente. To postižemo množenjem vrednosti iz matrice čistog toka sa vrednostima sopstvenih vektora. Dobijamo:

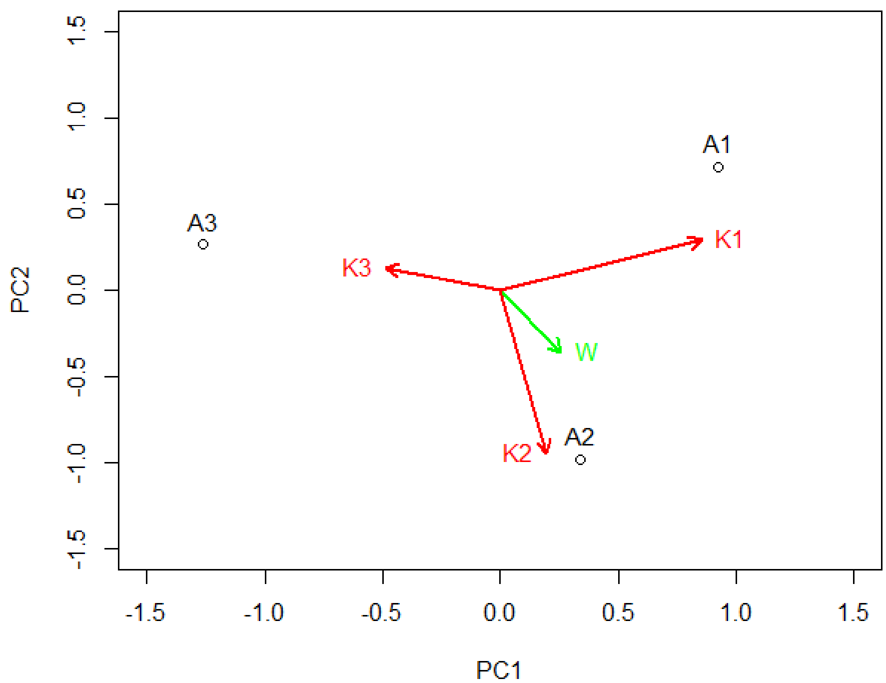
	K1	K2	K3
A1	1	-0,5	-0,356
A2	0	1	-0,289
A3	-1	-0,5	0,645

	GK1	GK2
K1	0,853	0,291
K2	0,197	-0,948
K3	-0,484	0,128

Ovim korakom dobijamo vrednosti alternativa u novom, dvodimenzionom prostoru. Projekciju kriterijuma na ovaj prostor dobijamo direktno iz matrice sopstvenih vektora. Vrednosti težina dobijamo na isti način, množenjem težina kriterijuma sa sopstvenim vrednostima.

GAIA ravan za naš primer izgleda kao na slici ispod. Alternative su prikazane crnim slovima (A1, A2 i A3), dok su kriterijumi prikazani crvenim linijama. Dužina crvene linije predstavlja uticaj kriterijuma na razdvajanje alternativa. Kriterijumi K1 i K2 jasno razdvajaju alternative, a K3 ih razdvaja, ali ne istim intenzitetom. Ukoliko su strelice sličnog smera, to je znak da su kriterijumi korelisani, što dalje znači da je moguće da kriterijumi mere istu stvar, te je potrebno detaljnije ispitati te kriterijume. Zelenom bojom obeležen je vector težina.

W	0,3 * 0,853 + 0,5 * 0,197 + 0,2 * (-0,484) = **0,2576**	0,3 * 0,291 + 0,5 * (-0,948) + 0,2 * 0,128 = **-0,3611**
	GK1	GK2
A1	1 * 0,853 + (-0,5) * 0,197 + (-0,356) * (-0,484) = **0,9269**	1 * 0,291 + (-0,5) * (-0,948) + (-0,356) * 0,128 = **0,7194**
A2	0 * 0,853 + 1 * 0,197 + (-0,289) * (-0,484) = **0,3369**	0 * 0,291 + 1 * (-0,948) + (-0,289) * 0,128 = **-0,985**
A3	(-1) * 0,853 + (-0,5) * 0,197 + 0,645 * (-0,484) = **-1,2637**	(-1) * 0,291 + (-0,5) * (-0,948) + 0,645 * 0,128 = **0,2656**



Najprihvatljivija alternativa je ona alternativa koja ima najveću projekciju na vektor težina. U ovom slučaju, to je alternativa A2. Postupno, do rešenja se dolazi na sledeći način:

W	0,2576	-0,3611	
	GK1	GK2	Očekivana korist
A1	0,9269	0,7194	0,9269 * 0,2576 + 0,7194 * (-0,3611) = -0,02104
A2	0,3369	-0,985	0,3369 * 0,2576 + (-0,985) * (-0,3611) = 0,44246
A3	-1,2637	0,2656	(-1,2637 * 0,2576) + 0,2656 *(-0,3611) = -0,42142