

# Domača naloga 2

Neža Kržan, Tom Rupnik

## Kazalo

<b>1</b>	<b>Opis podatkov</b>	<b>3</b>
<b>2</b>	<b>Korelacijska matrika</b>	<b>4</b>
<b>3</b>	<b>Metoda glavnih komponent</b>	<b>7</b>
3.1	Določitev števila komponent . . . . .	7
3.1.1	Podrobnejša analiza izbire števila glavnih komponent . . . . .	7
3.1.2	Paralelna analiza . . . . .	7
3.2	Uteži za izbrani glavni komponenti . . . . .	8
3.3	Delež nepojasnene variabilnosti . . . . .	9
3.4	Grafični prikaz glavnih komponent . . . . .	10
3.5	Zaključek . . . . .	11
<b>4</b>	<b>Faktorska analiza</b>	<b>12</b>
4.1	Določitev števila faktorjev . . . . .	12
4.2	Faktorska analiza . . . . .	14
4.3	Komunalitete . . . . .	15
4.4	Faktorske uteži - interpretacija skupnih faktorjev . . . . .	16
4.5	Grafični prikaz modela . . . . .	18
4.6	Zaključek . . . . .	19

## Slike

1	Prikaz Pearsonove korelacije med spremenljivkami. . . . .	5
2	Izbira števila glavnih komponent. . . . .	7
3	Scree diagram za izbiro števila glavnih komponent. . . . .	8
4	Grafični prikaz interpretacije metode glavnih komponent. . . . .	10
5	Grafični prikaz spremenljivk glede na komponento. . . . .	11
6	Scree diagram za izbiro števila faktorjev. . . . .	12
7	Izbira števila faktorjev. . . . .	13
8	Prikazi drugih pristopov za določanje št. faktorjev. . . . .	14
9	Matrika odklonov korelacij na podlagi modela od empiričnih korelacij(levo). . . . .	15
10	Grafični prikaz odklonov. . . . .	16
11	Grafični prikaz strukturnih uteži za pravokotno rotacijo na grafu. . . . .	16
12	Grafični prikaz modela. . . . .	18
13	Grafični prikaz razporeditve spremenljivk v faktorje. . . . .	19

## Tabele

1	Opisne statistike za številske spremenljivke v podatkovnem okviru. . . . .	4
3	Prikaz uteži glavnih komponent. . . . .	9

4	Komunalitete. . . . .	15
5	Prikaz faktorskih uteži. . . . .	17

# 1 Opis podatkov

Podatki so bili zbrani med letoma 2016-2018 preko interaktivnega testa osebnosti na spletu.

Test osebnosti je bil zgrajen na podlagi "Big-Five Factor Markers" pridobljivih iz IPIP (International Personality Item Pool).

Lestvica, uporabljena za odgovore, je bila od 1 do 5 oziroma od -1 do -5 (podrobnejše razlage ni bilo dostopne, kot 1 = se ne strinjam, 3 = nevtrarno, 5 = se strinjam).

Za domačo nalogo sva si izbrala spremenljivke, ki so bile kodirane z "EXT1-10" in "EST1-10". Te oznake predstavljajo trditev oz. osebnostno lastnost, katero so anketiranci ovrednotili po zgornji lestvici. Pomen, pri čemer je s + oz. - označeno ali je trditev bila označena kot pozitivna (velja za ta sklop/lastnost), ali negativno (ne velja za ta sklop/lastnost, lestvica je obrnjena):

## Energija (ekstravertnost, surgentnost)

- EXT1...Sem zabavna oseba, ki je rada v središču dogajanja (+)
- EXT2...Ne govorim veliko (-)
- EXT3...Dobro se počutim v krogu ljudi (+)
- EXT4...Raje sem v ozadju (stran od pozornosti/neopazen) (-)
- EXT5...Jaz začenjam pogovore (+)
- EXT6...Nimam veliko za povedati/Sem redkobeseden (-)
- EXT7...Na zabavah se pogovarjam z veliko različnimi osebami (+)
- EXT8...Ne maram pritegniti pozornosti ostalih (-)
- EXT9...Ne moti me če sem v središču dogajanja (+)
- EXT10...V okolici neznancev ne govorim (-)

## Čustvena stabilnost (nevroticizem)

- EST1...Hitro postanem živčen (-)
- EST2...Večino časa sem sproščen (+)
- EST3...Skrbi me za veliko stvari (-)
- EST4...Redko sem žalosten, melanholičen, emotično pretresen (+)
- EST5...Stvari me hitro zmotijo (-)
- EST6...Hitro se razjezim (-)
- EST7...Dosti spreminjam razpoloženje (-)
- EST8...Imam pogosta nihanja razpoloženja (-)
- EST9...Sem hitro vzkipljiv (-)
- EST10...Pogosto se počutim žalostno, melanholično, emotično pretreseno (-)

Vrednosti trditev, ki so označene z (-) bomo množili z  $-1$  in s tem obrnili lestvico.

Iz podatkovnega okvirja sva odstranila vrstice, ki so vsebovale kakšno *NA* vrednost. Takih vrstic je bilo 280, torej imava na koncu 4841 podatkov.

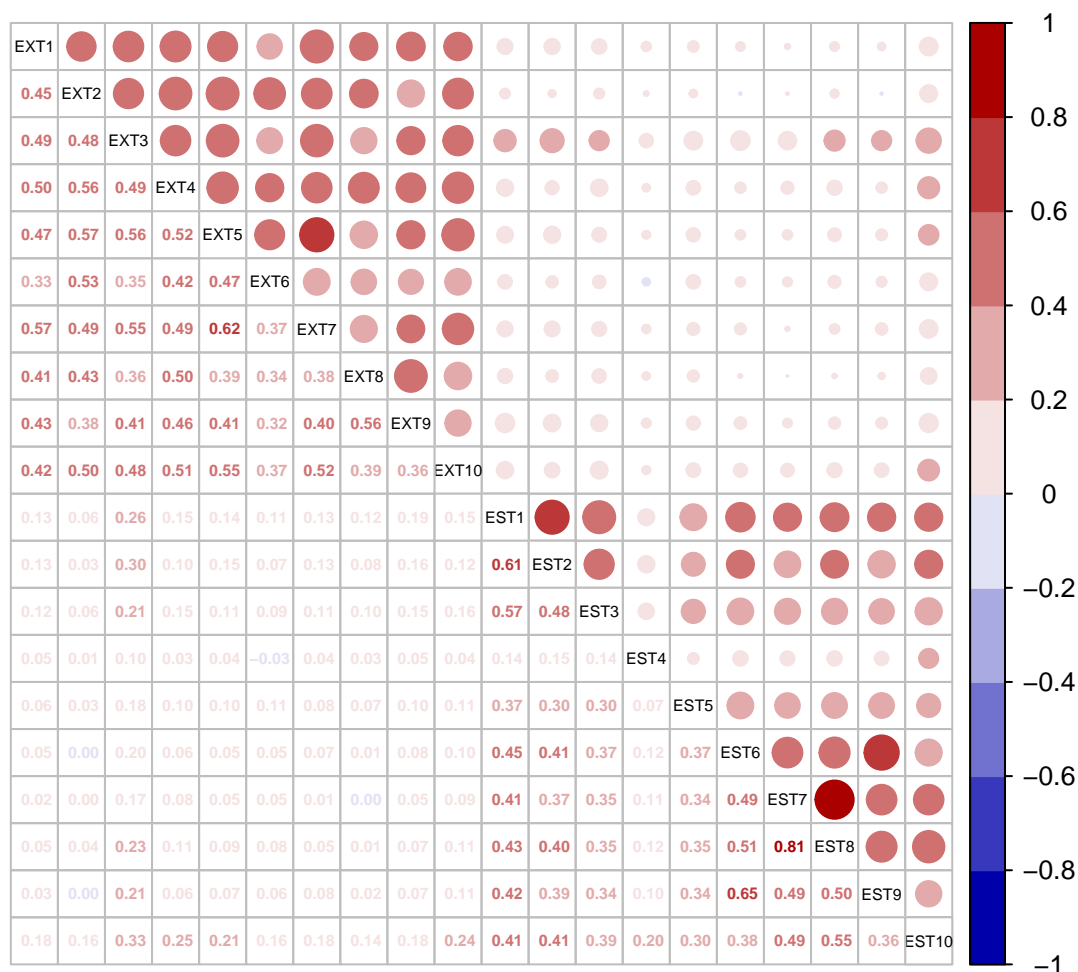
Poglejmo si nekatere opisne statistike sklopa spremenljivk pri *energiji* in sklopa spremenljivk *čustvene stabilnosti*. V spodnji tabeli vidimo, da so podatki v redu in nimamo nepravilnih vrednosti, ki niso znotraj lestvice.

Tabela 1: Opisne statistike za številske spremenljivke v podatkovnem okviru.

spremenljivke	mean	Min	Max
EXT1	2.480	1	5
EXT2	-2.972	-5	-1
EXT3	3.055	1	5
EXT4	-3.242	-5	-1
EXT5	3.118	1	5
EXT6	-2.469	-5	-1
EXT7	2.679	1	5
EXT8	-3.414	-5	-1
EXT9	2.944	1	5
EXT10	-3.688	-5	-1
EST1	-3.325	-5	-1
EST2	3.061	1	5
EST3	-3.862	-5	-1
EST4	2.804	1	5
EST5	-3.182	-5	-1
EST6	-2.913	-5	-1
EST7	-3.107	-5	-1
EST8	-2.884	-5	-1
EST9	-3.015	-5	-1
EST10	-3.122	-5	-1

## 2 Korelacijska matrika

Ker želimo pojasniti čim več variabilnosti spremenljivk in korelacijo med spremenljivkami, se bomo to lotili z metodo glavnih komponent in faktorsko analizo. Metodi sta smiselni v primeru, ko so spremenljivke v podatkovnem okviru med seboj korelirane, zato si pogledjmo korelacijsko matriko.

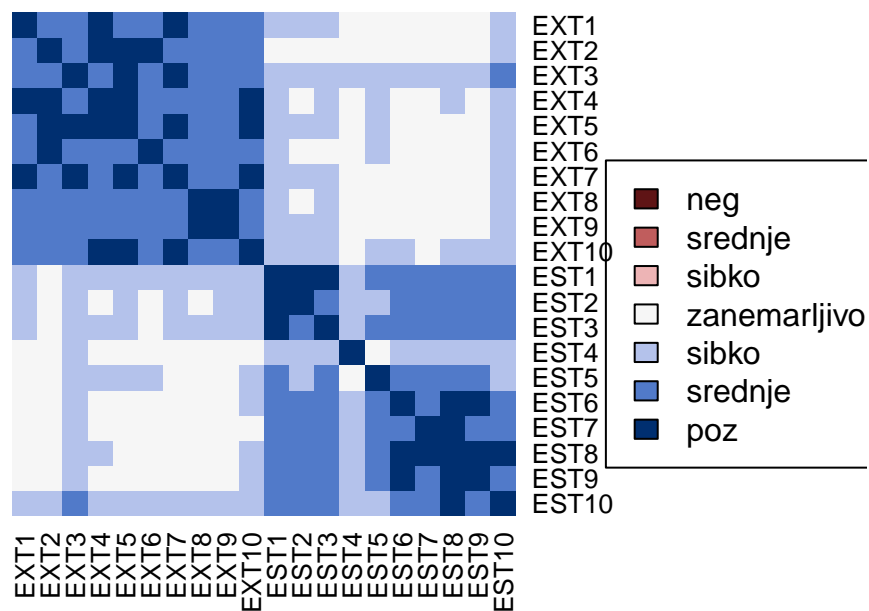


Slika 1: Prikaz Pearsonove korelacije med spremenljivkami.

Iz matrike korelacije vidimo, da se korelacija znotraj spremenljivk **EXT** giblje približno med 0.3-0.6. Pri spremenljivkah **EST** je razpon vrednosti nekoliko večji in sicer 0.1-0.8. Pri tem z nizkimi vrednostmi izstopa predvsem spremenljivka **EST4**, ki predstavlja trditev "Redko sem žalosten, melanholičen, emotično pretresen".

Opaziti je, da so korelacije znotraj skupin spremenljivk (EXT in EST) večje, zato lahko na spodnjem prikazu pričakujemo dva očitna sklopa.

Korelacije znotraj obeh sklopov so pozitivne in srednje oziroma šibke, edino povezanost med **EST8** in **EST7** je močnejša. Spremenljivke med sklopoma pa so zanemarljivo povezane in se jasno vidi razlika med sklopoma.



Res vidimo, da sta tvorjeni dve skupini spremenljivk, in sicer **EXT1-10** in **EST1-10**.

Da se zares prepričamo, da spremenljivke med seboj niso neodvisne lahko izvedemo Bartlettov test, ki ima ničelno hipotezo “*vse spremenljivke so neodvisne*”, kar bi pomenilo, da je korelacijska matrika spremenljivk identiteta. Ko izvedemo test dobimo p-vrednost enako 0, kar pomeni, da ovržemo ničelno hipotezo. Torej spremenljivke med seboj niso neodvisne oz. so korelirane in metodi, ki jih bomo izvajali v nadaljevanju, sta smiselni.

Zanima nas še, ali bomo imeli kakšne težave pri faktorski analizi, zato izvedemo še Kaiser-Meyer-Olkin test.

Test nam vrne naslednje vrednosti:

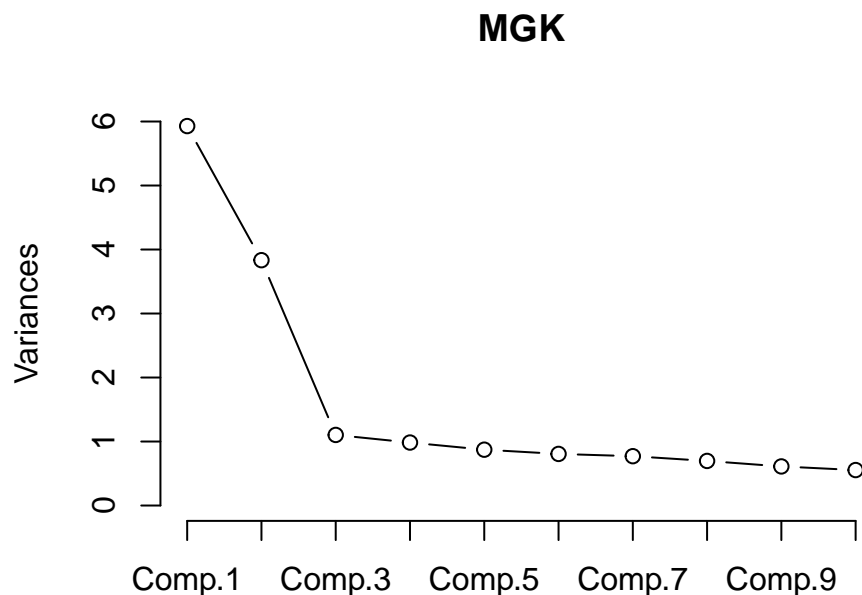
EXT1	EXT2	EXT3	EXT4	EXT5	EXT6	EXT7	EXT8	EXT9	EXT10
0.93	0.92	0.95	0.94	0.93	0.92	0.92	0.89	0.91	0.95
EST1	EST2	EST3	EST4	EST5	EST6	EST7	EST8	EST9	EST10
0.88	0.89	0.91	0.87	0.96	0.88	0.80	0.80	0.87	0.94

Vse vrednosti so večje oz. enake 0.8, kar pomeni da so podatki primerni za faktorsko analizo in ne bomo imeli težav.

### 3 Metoda glavnih komponent

Cilj metode glavnih komponent je s prvimi nekaj nekoreliranimi glavnimi komponentami pojasniti kar največ variabilnosti originalnih spremenljivk in zmanjšati razsežnost podatkov tako, da se izgubi čim manj informacij. Glavne komponente so linearne kombinacije merjenih spremenljivk in so med seboj neodvisne. Urejene so glede na pojasnenost variabilnosti od največ do najmanj.

#### 3.1 Določitev števila komponent



Slika 2: Izbira števila glavnih komponent.

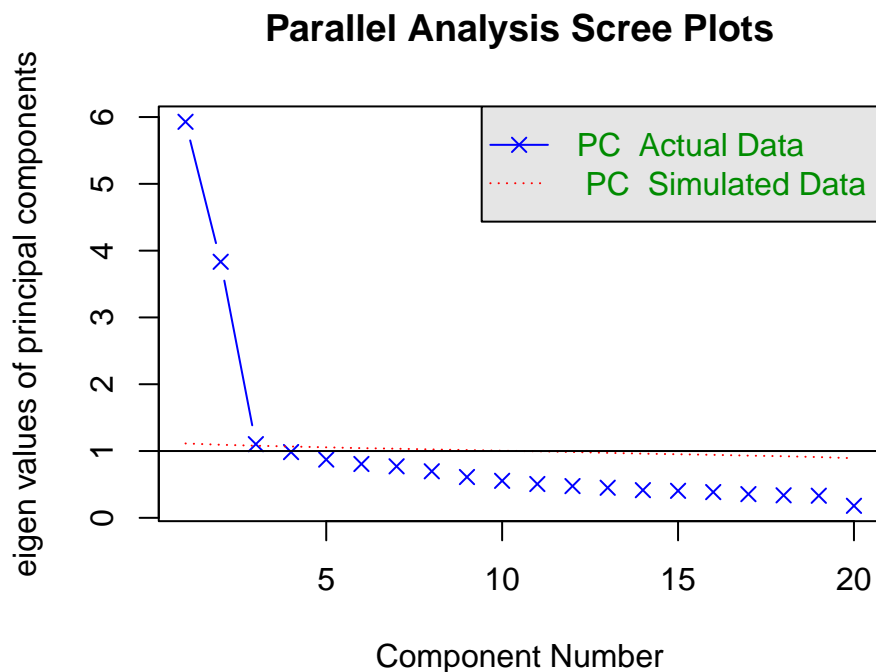
Prikazan je graf variance v odvisnosti od števila komponent (izvedeno na standardiziranih podatkih). Izbrano število glavnih komponent je tam, kjer se graf lomi oziroma ima koleno in so vrednosti nad 1, pri čemer obdržimo komponente nad kolenom. Zato bi se odločili za 2 komponenti. Tudi vrednost pri 3 komponentah malo preseže vrednost nad 1, vendar je koleno pri 3 komponentah bolj očitno zato se vseeno raje odločimo za 2 komponenti.

##### 3.1.1 Podrobnejša analiza izbire števila glavnih komponent

Ponavadi se odločimo za tisto število komponent, da so lastne vrednosti komponent večje od povprečne variabilnosti standardiziranih spremenljivk, ki je enaka 1 (oz. standardni odklon večji od 1). pri nas to velja za 3 komponente. Zadnja je sicer ponovno zelo na meji, zato se zaenkrat raje odločimo za dve komponenti.

##### 3.1.2 Paralelna analiza

Glede števila glavnih komponent se lahko odločimo tudi na podlagi scree diagrama, ki grafično predstavi lastne vrednosti. Izbrano število glavnih komponent je ponovno tam, kjer se graf lomi oziroma ima koleno, pri čemer obdržimo komponente nad kolenom.



Slika 3: Scree diagram za izbiro število glavnih komponent.

S paralelno analizo smo dobili nekoliko drugačen rezultat kot v prejšnjih primerih. Predlagano število glavnih komponent je 3 vendar je ta vrednost na meji (je tudi težko razvidno iz grafa), zato bomo obržali število glavnih komponent 2.

Torej, na podlagi paralelne analize, primerjav lastnih vrednosti, ki morajo biti večje od 1 in na podlagi primerjav lastnih vrednosti se odločimo za 2 glavni komponenti, kljub temi, da pri vseh treh načinih ni popolnoma razvidna meja med dvema ali tremi komponentami.

### 3.2 Uteži za izbrani glavni komponenti

Uteži dobljenih glavnih komponent reskaliramo, ker nereskalirane uteži predstavljajo regresijske koeficiente. Reskalirana utež pa predstavljajo korelacije med komponentami in spremenljivkami in s pomočjo teh korelacij lahko interpretiramo glavne komponente.



Tabela 3: Prikaz uteži glavnih komponent.

	Nereskalirane uteži		Reskalirane uteži	
	Comp.1	Comp.2	Comp.1	Comp.2
EXT1	0.2450903	0.1999064	0.5967717	0.3913866
EXT2	0.2436205	0.2475418	0.5931927	0.4846495
EXT3	0.2977556	0.0980326	0.7250066	0.1919332
EXT4	0.2718738	0.1957218	0.6619869	0.3831938
EXT5	0.2750991	0.2087335	0.6698403	0.4086688
EXT6	0.2141123	0.1693563	0.5213431	0.3315742
EXT7	0.2626848	0.2101527	0.6396125	0.4114474
EXT8	0.2189377	0.1989599	0.5330926	0.3895336
EXT9	0.2369629	0.1578704	0.5769822	0.3090865
EXT10	0.2613589	0.1657963	0.6363841	0.3246041
EST1	0.2216208	-0.2563855	0.5396256	-0.5019642
EST2	0.2044839	-0.2428376	0.4978990	-0.4754395
EST3	0.1955001	-0.2210849	0.4760243	-0.4328510
EST4	0.0673235	-0.0835973	0.1639262	-0.1636710
EST5	0.1589855	-0.2004730	0.3871146	-0.3924959
EST6	0.1826405	-0.3043970	0.4447124	-0.5959635
EST7	0.1820530	-0.3229927	0.4432817	-0.6323710
EST8	0.2025200	-0.3166841	0.4931170	-0.6200199
EST9	0.1788389	-0.2927878	0.4354558	-0.5732345
EST10	0.2379673	-0.1966613	0.5794278	-0.3850332

Interpretacija reskaliranih uteži je, da so spremenljivke čustvene stabilnosti negativno korelirane z drugo glavno komponento, spremenljivke energije pa so močnejše pozitivno povezane s prvo glavno komponento kot z drugo.

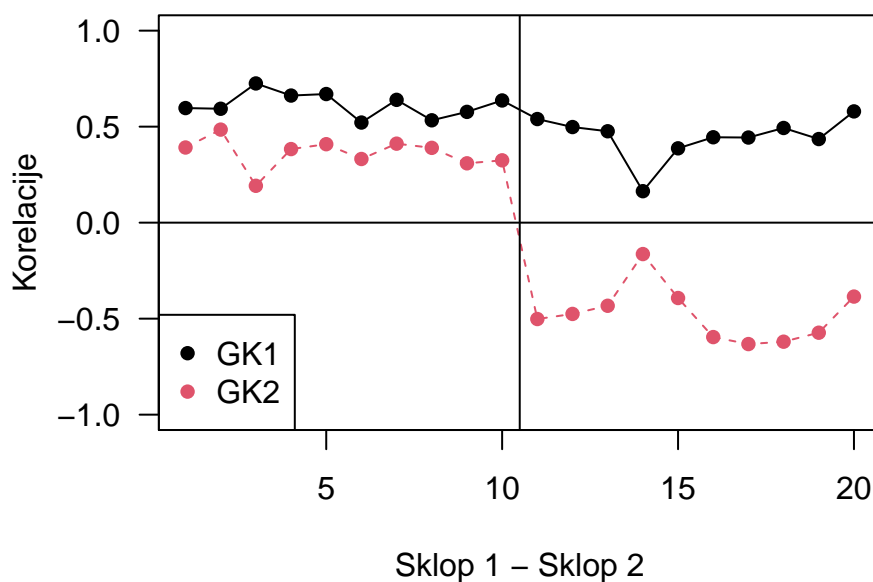
Če strnemo, prva glavna komponenta povzema predvsem pomembnost energije in druga glavna komponenta povzema predvsem pomembnost čustvene stabilnosti (negativne korelacije).

### 3.3 Delež nepojasnjene variabilnosti

Velja, da je vsota kvadratov reskaliranih uteži po spremenljivkah varianca spremenljivke. Zanima nas celotna variabilnost, torej koliko prispeva vsaka komponenta.

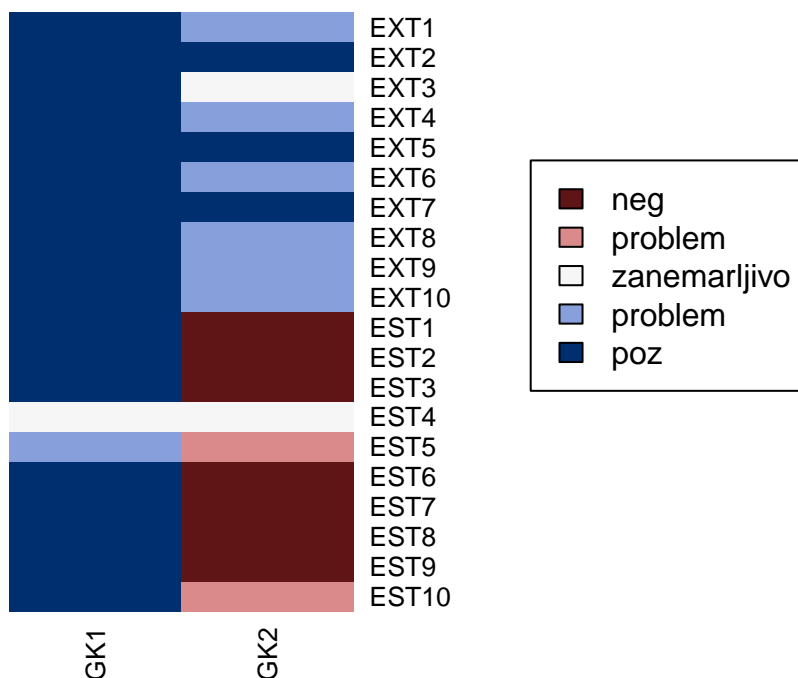
Pri nas s prvo komponento pojasnimo 29.6% celotne variabilnosti in z drugo komponento 19.1%, skupaj pa 48.8% celotne variabilnosti. Z izbranimi prvima glavnima komponentama pojasnimo največji delež EST8 (63%), EST7 (60%) in EXT2 (59%). Najmanjši delež variabilnosti pa pojasnimo pri spremenljivki EST4 (5%).

### 3.4 Grafični prikaz glavnih komponent



Slika 4: Grafični prikaz interpretacije metoe glavnih komponent.

Z navpično črto na grafu smo ločili sklopa spremenljivk in vidimo, da se komponenti ne obračata glede na navpično črto. Če so uteži nad 0.4 (vrednost korelacije) je v redu, kjer je korelacija slaba pa si želimo pod 0.2. Obe glavni komponenti sta v levi polovici pozitivno korelirani, ampak ima GK1 višje vrednosti (nad 0.5), GK2 ima pa vrednosti korelacije med 0.2-0.4, kar je območje, ki nam povzroča problem, saj korelacija ni zanemarljiva, ampak ni tudi tako visoka, da bi jo z zagotovostjo povezali s temi lastnostmi. Na desni polovici pa se glavni komponenti lepo ločita na pozitivno in negativno koreliranost in so vrednosti po absolutni vrednosti nad 0.4 z nekaj izjemami (EST4, EST5).



Slika 5: Grafični prikaz spremenljivk glede na komponento.

Problematična je spremenljivka **EST4** saj ne spada nikamor, kar smo videli že zgoraj pri nizki korelaciji. Ni jasno ločenih spremenljivk glede na komponento.

### 3.5 Zaključek

Pri tej metodi sva se odličila za dve glavni komponenti, sicer je bilo skoraj vedno na meji, da bi bilo bolje izbrati tri. Odločila sva se za dve in sicer spremenljivke, ki ocenjujejo energijo so v eni in spremenljivke, ki ocenjujejo čustveno stabilnost v drugi. Z obema komponentama smo pojasnili skoraj 50% celotne variabilnosti, pojasnili pa smo največji delež spremenljivke EST8 (63%), najmanjši delež variabilnosti pa spremenljivke EST4 (5%).

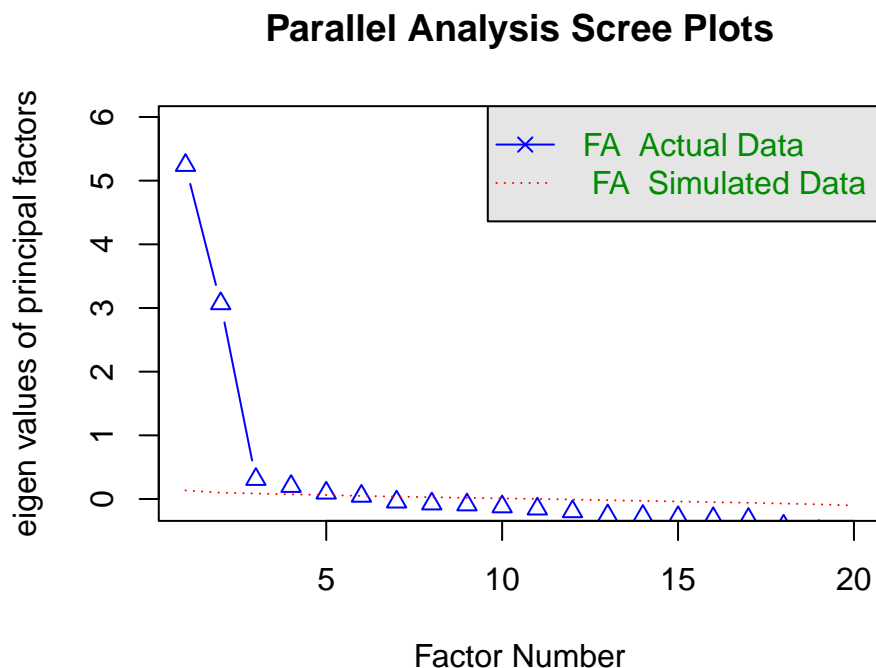
## 4 Faktorska analiza

Faktorska analiza poizkuša poenostaviti kompleksnost povezav med opazovanimi spremenljivkami tako, da poskuša najti novo množico spremenljivk, ki predstavlja skupno opazovanim spremenljivkam. Ponavadi jo uporabimo takrat, ko pojmov ne moremo meriti neposredno in zato izberemo nekaj direktno merjenih spremenljivk, ki so indikatorji tistih, ki jih želimo meriti. Preko faktorjev poskušamo čim bolj pojasniti korelacije oziroma kovariance med merjenimi spremenljivkami. Merjene spremenljivke so tako lineane kombinacije skupnih in specifičnih faktorjev.

### 4.1 Določitev števila faktorjev

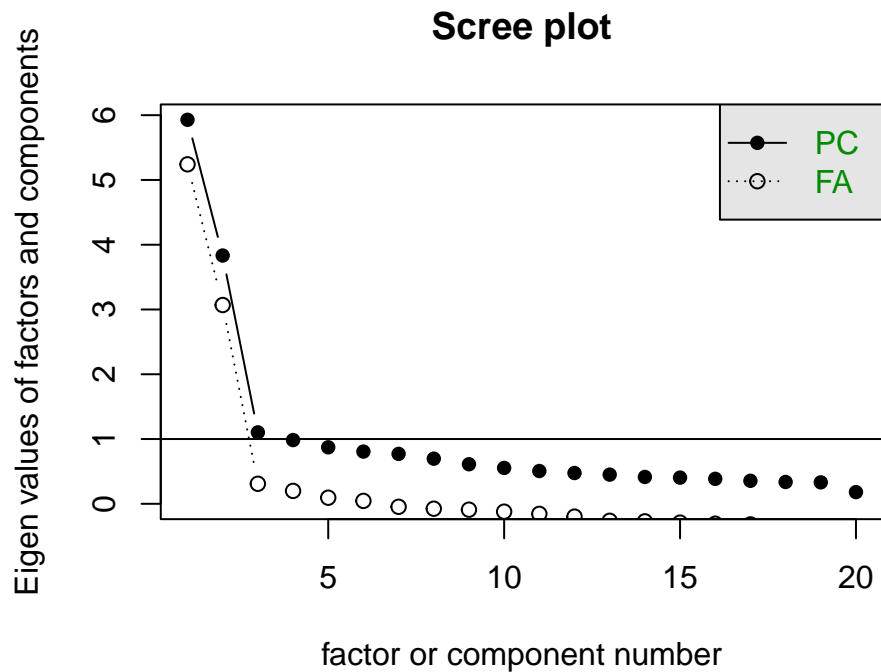
Za določitev števila faktorjev uporabljamo podobne kriterije kot pri metodi glavnih komponent. Odločimo se lahko na podlago kolena scree diagrama, na podlagi lastnih vrednosti faktorjev ali pa s primerjavo lastnih vrednosti faktorjev s tistimi na podlagi podatkov generiranih iz nekoreliranih spremenljivk.

Predpostaviti moramo, da se spremenljivke porazdeljujejo normalno in tako lahko uporabimo faktorsko metodo največjega verjetja.



Slika 6: Scree diagram za izbiro števila faktorjev.

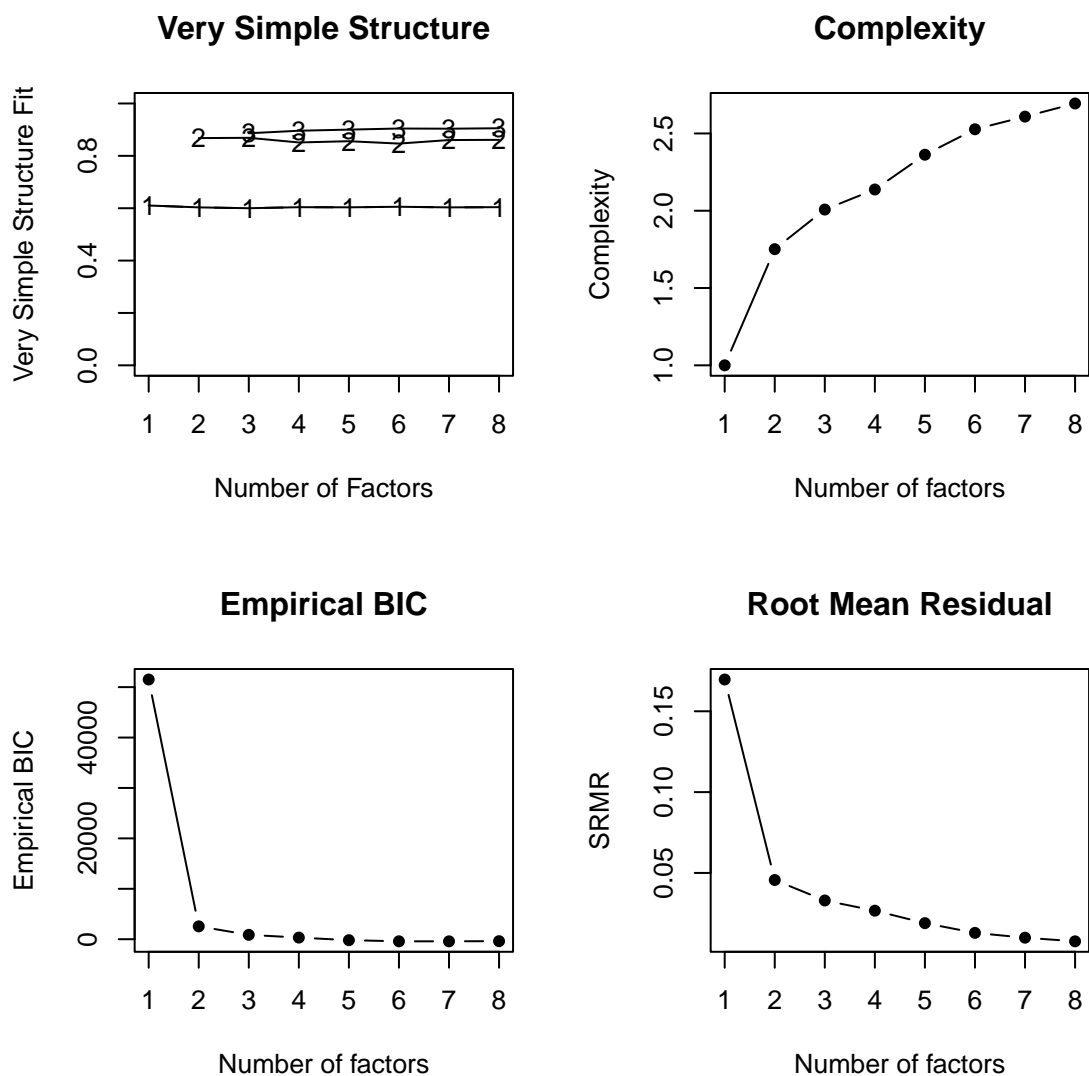
Pri paralelni analizi gledamo koliko vrednosti je nad rdečo črto in zelo očitno je, da bi se odločili za dva faktorja oziroma tri. Analiza pa sicer predlaga pet faktorjev, ker gleda koliko vrednosti je večjih od simuliranih podatkov.



Slika 7: Izbira števila faktorjev.

Podobno kot pri metodi glavnih komponent, si tudi tukaj lahko pomagamo s screeplotom, kjer gledamo kolena in koliko komponent je večjih od 1 za FA vrednosti. Na podlagi zgornjega grafa je zelo očitno, da bi lahko imeli dva faktorja.

Poglejmo si še rezultat pristopa VSS in eBIC. Pri VSS izberemo število, kjer so uteži pri spremenljivkah največje in s tem največ pojasnijo (upoštevane so samo najvišje uteži pri spremenljivkah). Pristop eBIC pa število faktorjev določi tam, kjer je koleno.



Slika 8: Prikazi drugih pristopov za določanje št. faktorjev.

Na podlagi VSS predlaga 3 faktorje, pri čemer upoštevamo uteži 1 in 2, saj po treh faktorjih oz. že po dveh zelo počasi narašča. eBIC minimum doseže pri 7 faktorjih, vendar je najboljše število faktorjev dva, saj prileganje pri večjem številu faktorjev postaja zanemarljivo boljše.

Po vseh sklepih se odločimo, da obdržimo dva faktorja.

## 4.2 Faktorska analiza

Na podlagi izbranega števila faktorjev tj. dveh faktorjev bomo izvedli faktorsko analizo.

Najprej začnemo s poševno rotacijo (oblimin) vidimo, da je korelacija med faktorjema zelo majhna (0.188) in manjša od 0.2. Zato naredimo še pravokotno rotacijo (varimax), kjer so korelacije med faktorji enake 0.

Pri pravokotni rotaciji si želimo, da bi bil indeks Tucker Lewis večji od 0.95 ali RMSE manjša od 0.05, da je faktorska analiza zares dobra. Naš indeks je 0.82, RMSE pa 0.091, kar kaže na to, da faktorski model ni najboljši.

S prvim faktorjem pojasnimo 23% z drugim pa 20% (torej z dvema faktorjema pojasnimo 43%) variabilnosti med spremenljivkami, izmed vse pojasnene variabilnosti z izbranimi dvema faktorjema s prvim faktorjem

pojasnimo 54% variabilnosti, z drugim pa 46%.

### 4.3 Komunalitete

Zanimajo nas komunalitete, ki odražajo del variabilnosti spremenljivke, ki je pojasnjen s skupnimi faktorji. Z izbranimi skupnimi faktorji smo pojasnili naslednje deleže posameznega indikatorja:

Tabela 4: Komunalitete.

EST4	EST5	EST3	EXT6	EXT9
0.0343038	0.2284565	0.3044014	0.3225001	0.3559597

EXT8	EST2	EST1	EST10	EST9
0.3573577	0.3585661	0.4175793	0.4364631	0.4399598

EXT1	EXT10	EST6	EXT3	EXT4
0.4540566	0.4644312	0.465118	0.5269098	0.5288192

EXT2	EXT7	EXT5	EST7	EST8
0.5288659	0.5438011	0.5847232	0.6410519	0.6745271

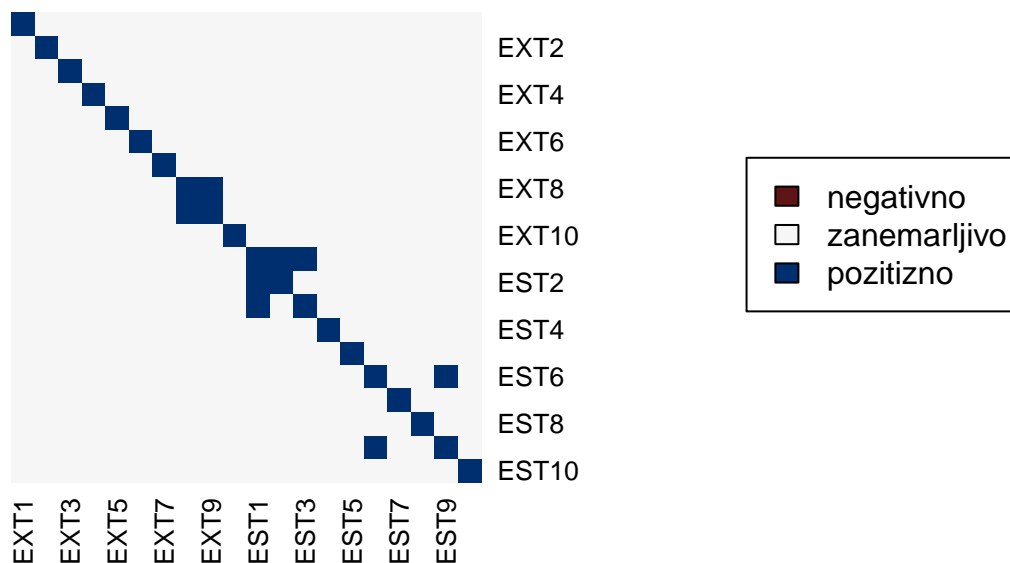
V našem primeru je večina komunalitet nad 0.2, kar pomeni, da je spremenljivka dobro pojasnjena s skupnim faktorskim modelom. Odstopa zgolj **EST4**, ki ima vrednost komunalitet nižjo od 0.2 in je njena variabilnost slabše pojasnjena z modelom.

Na spodnjih grafih vidimo matriko odklonov korelacij na podlagi modela od empiričnih korelacij. Izven diagonale so razlike ocenjenih in empiričnih korelacij, na diagonali pa so unikvitete(angl. *uniquenesses*), ki povedo kolikšen delež variabilnosti spremenljivke ni pojasnjen s skupnimi faktorji.

	EXT1	EXT2	EXT3	EXT4	EXT5	EXT6	EXT7	EXT8	EXT9	EXT10	EST1	EST2	EST3	EST4	EST5	EST6	EST7	EST8	EST9	EST10
EXT1	55	-4	2	1	-4	-5	8	1	3	-3	0	2	1	1	-1	0	0	-1	-1	-1
EXT2	-4	47	-1	4	2	12	-4	-1	-4	1	-4	-6	-3	-1	-2	-1	3	3	-1	0
EXT3	2	-1	47	-3	2	-5	4	-5	-1	0	-1	6	-2	3	0	0	-2	-1	2	1
EXT4	1	4	-3	47	-4	0	-5	7	3	2	-1	-5	1	-1	0	-2	2	2	-2	2
EXT5	-4	2	2	-4	42	4	5	-6	-4	3	-2	1	-3	-1	0	-1	1	1	0	-1
EXT6	-5	12	-5	0	4	68	-5	0	-1	-2	-1	-3	-1	-7	3	0	1	2	1	-1
EXT7	8	-4	4	-5	5	-5	46	-6	-4	2	-1	0	-1	0	0	2	-1	-1	3	-3
EXT8	1	-1	-5	7	-6	0	-6	64	21	-1	2	-1	2	1	2	-1	0	-1	0	-1
EXT9	3	-4	-1	3	-4	-1	-4	21	64	-5	5	3	2	1	1	1	-1	-2	-1	-1
EXT10	-3	1	0	2	3	-2	2	-1	-5	54	-2	-4	1	-1	-1	1	1	0	1	1
EST1	0	-4	-1	-1	-2	-1	-1	2	5	-2	58	22	21	3	6	2	-9	-9	-1	-1
EST2	2	-6	6	-5	1	-3	0	-1	3	-4	22	64	15	4	1	1	-10	-8	0	2
EST3	1	-3	-2	1	-3	-1	-1	2	2	1	21	15	70	3	4	0	-7	-9	-2	2
EST4	1	-1	3	-1	-1	-7	0	1	1	-1	3	4	3	97	-2	0	-3	-3	-2	8
EST5	-1	-2	0	0	0	3	0	2	1	-1	6	1	4	-2	77	5	-4	-4	3	-1
EST6	0	-1	0	-2	-1	0	2	-1	1	1	2	1	0	0	5	53	-5	-5	20	-5
EST7	0	3	-2	2	1	1	-1	0	-1	1	-9	-10	-7	-3	-4	-5	36	16	-4	0
EST8	-1	3	-1	2	1	2	-1	-1	-2	0	-9	-8	-9	-3	-4	-5	16	33	-4	4
EST9	-1	-1	2	-2	0	1	3	0	-1	1	-1	0	-2	-2	3	20	-4	-4	56	-6
EST10	-1	0	1	2	-1	-1	-3	-1	-1	1	-1	2	2	8	-1	-5	0	4	-6	56

\* all values in cells were multiplied by 100

Slika 9: Matrika odklonov korelacij na podlagi modela od empiričnih korelacij(levo).



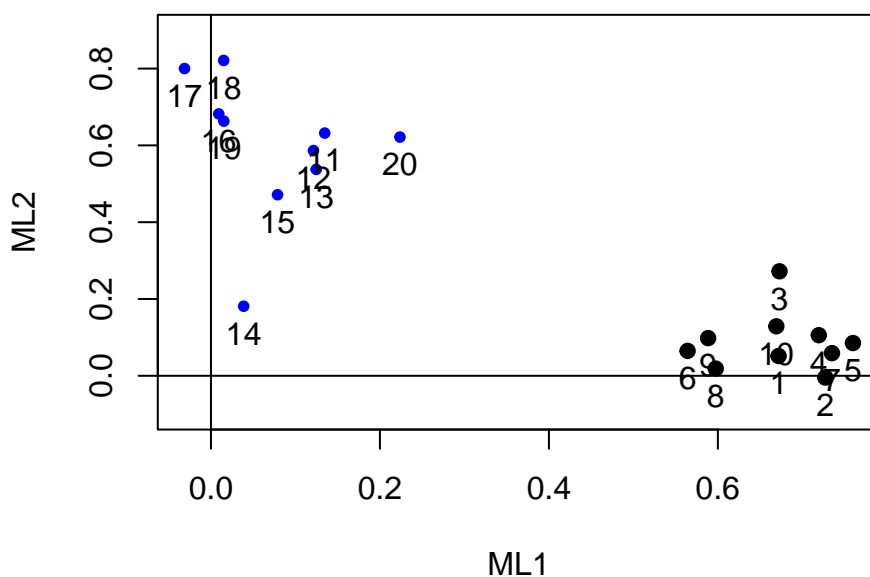
Slika 10: Grafični prikaz odklonov.

Vidimo, da so unikvitete visoke, ampak manjše od 0.8, kar si želimo, z izjemo ene **EST4** - to smo že videli pri komunalitetah zgoraj, da ni dobro pojasnena s skupnim faktorjskim modelom. Pri ostalih spremenljivkah so vrednosti unikvitet pod 0.8 in njihove variabilnosti so tako dobro pojasnjene. Izven diagonal so vrednosti po absolutni vrednosti v večini blizu 0, kar je dobro, izrazito odstopajo le spremenljivke EXT8 in EXT9, EST1, EST2, EST3 in EST6, EST9.

#### 4.4 Faktorske uteži - interpretacija skupnih faktorjev

Pokažimo regresijske koeficiente ter strukturne uteži oziroma korelacijske koeficiente, ki so za pravokotno (varimax) rotacijo, ki sva jo v najinem primeru upoštevala.

#### Factor Analysis



Slika 11: Grafični prikaz strukturnih uteži za pravokotno rotacijo na grafu.



Želimo si situacijo, ko je en sklop uteži desno na x-osi spodnjega grafa in en sklop proti vrhu y-osi spodnjega grafa. Na x-osi se sklop uteži zelo lepo vidi, s tem ko na y-osi ni tako zelo lepo.

Tabela 5: Prikaz faktorskih uteži.

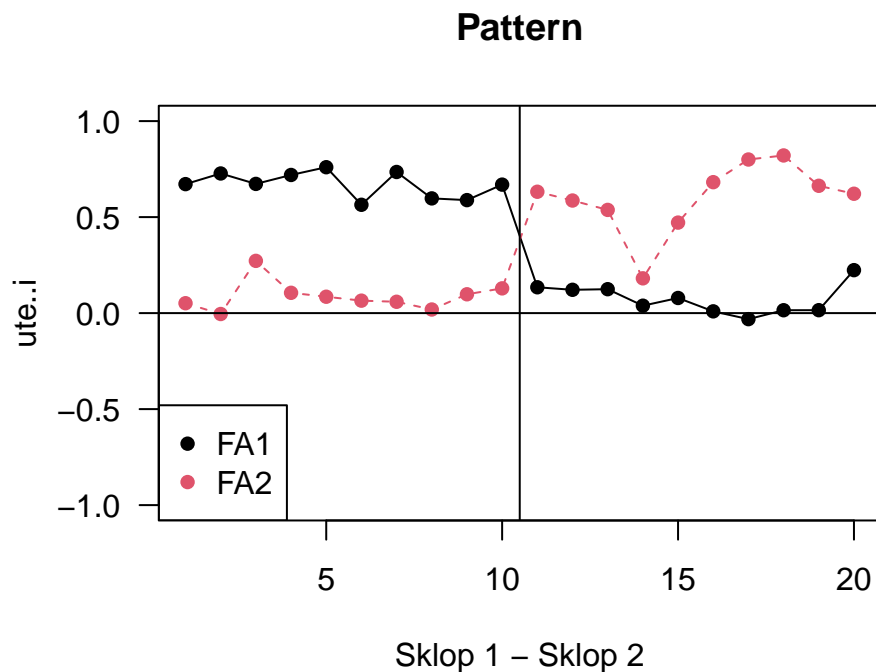
	Faktor1	Faktor2
EXT1	0.6718663	0.0515003
EXT2	0.7272166	-0.0046806
EXT3	0.6729541	0.2721078
EXT4	0.7194764	0.1057021
EXT5	0.7599037	0.0852620
EXT6	0.5641751	0.0648582
EXT7	0.7350679	0.0589598
EXT8	0.5974992	0.0187726
EXT9	0.5885044	0.0980929
EXT10	0.6691914	0.1288957
EST1	0.1346351	0.6320227
EST2	0.1213335	0.5863823
EST3	0.1245825	0.5374762
EST4	0.0386768	0.1811294
EST5	0.0788070	0.4714297
EST6	0.0092001	0.6819335
EST7	-0.0313368	0.8000437
EST8	0.0150238	0.8211586
EST9	0.0152286	0.6631198
EST10	0.2236806	0.6216351

Če si najprej pogledamo prvi faktor, vidimo, da ima visoke pozitivne, srednje do močne, korelacije s sklopom spremenljivk, ki meri energijo(EXT) in nizke korelacije s spremenljivkami, ki merijo čustveno stabilnost(EST). Izstopa le EST10, ki ima nekoliko višjo korelacijo, kot druge spremenljivke v sklopu čustvene stabilnosti.

Drugi faktor ima visoke pozitivne srednje do močne korelacije s spremenljivkami, ki merijo čustveno stabilnost in nižjo oz. šibko korelacijo s spremenljivkami, ki merijo energijo. Pri sklopu čustvene stabilnosti izstopa le z nižjo korelacijo spremenljivka EST4, pri sklopu energija pa ima višjo korelacijo od ostalih spremenljivka EXT3.

## 4.5 Grafični prikaz modela

Poglejmo si še grafični prikaz modela.

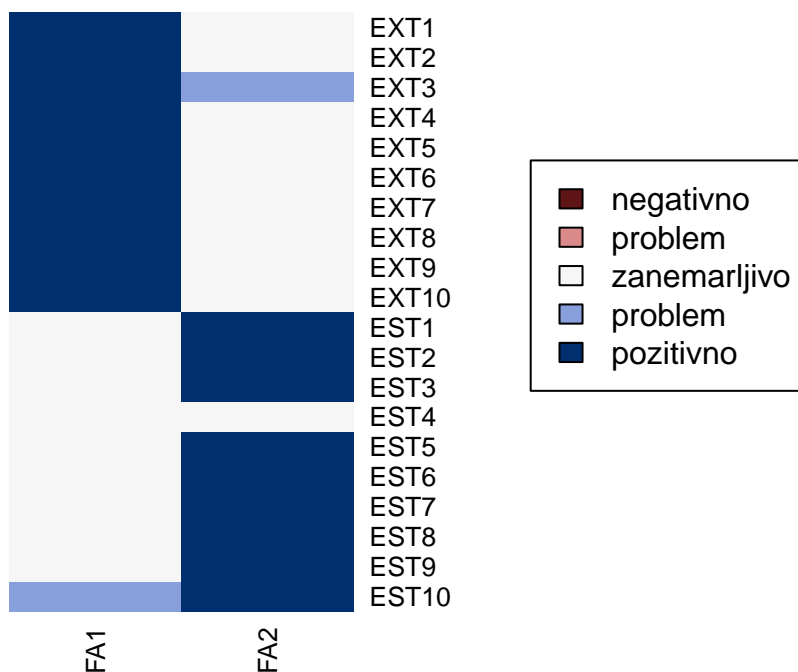


Slika 12: Grafični prikaz modela.

Z navpično črto na grafu smo ponovno ločili sklopa spremenljivk in vidimo, da se faktorja obračata glede na navpično črto, kar si želimo. V enem sklopu so uteži, ki so bolj korelirne višje od 0.4 in tiste, ki so manj so pod 0.2, z izjemo ene, ki je v prvem sklopu nekoliko višja in v drugem nekoliko nižja pri drugem faktorju. Pri faktorju ena pa so uteži zelo lepo urejene, nad 0.4 in pod 0.2.

Vidimo lahko, da je prvi faktor res lepo urejen, z izjemo ene spremenljivke, ki se sili v njega(EST10); to je bolje vidno na spodnjem grafu.

Spremenljivka EST4 ni v nobene faktorju, kar smo videli že prej, da je njena variabilnost slabše pojasnena z modelom. V drugi faktor se sili spremenljivka EXT3, ki naj bi bila v prvem faktorju.



Slika 13: Grafični prikaz razporeditve spremenljivk v faktorje.

## 4.6 Zaključek

Tukaj smo se odločili za dva faktorja in pravokotno rotacijo. Spremenljivke, ki so v povezavi z energijo so bile razdeljene v en faktorj, spremenljivke čustvene stabilnosti pa v drug. S prvim faktorjem smo pojasnili 23% in z drugim 43% variabilnosti med spremenljivkami. Pri tej metodi, so se spremenljivke bolje ločevale glede na faktor, težave smo imeli le pri spremenljivki EST4, ki ni bila dobro pojasnena s skupnim faktorskim modelom. Najbolje izmed vseh pa smo s faktorskim modelom pojasnili spremenljivko EST8.