



Metoda glavnih komponent

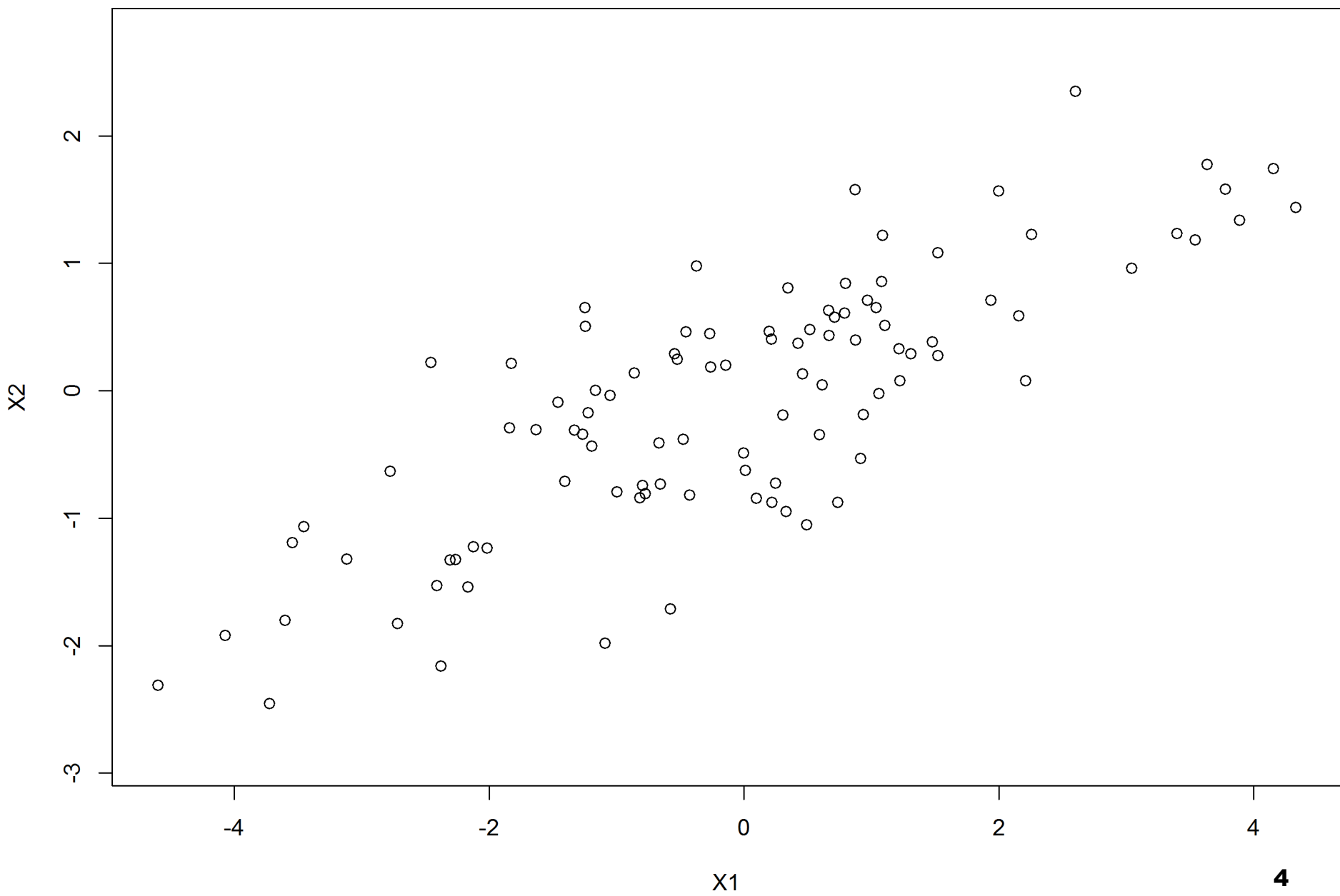
Multivariatna analiza

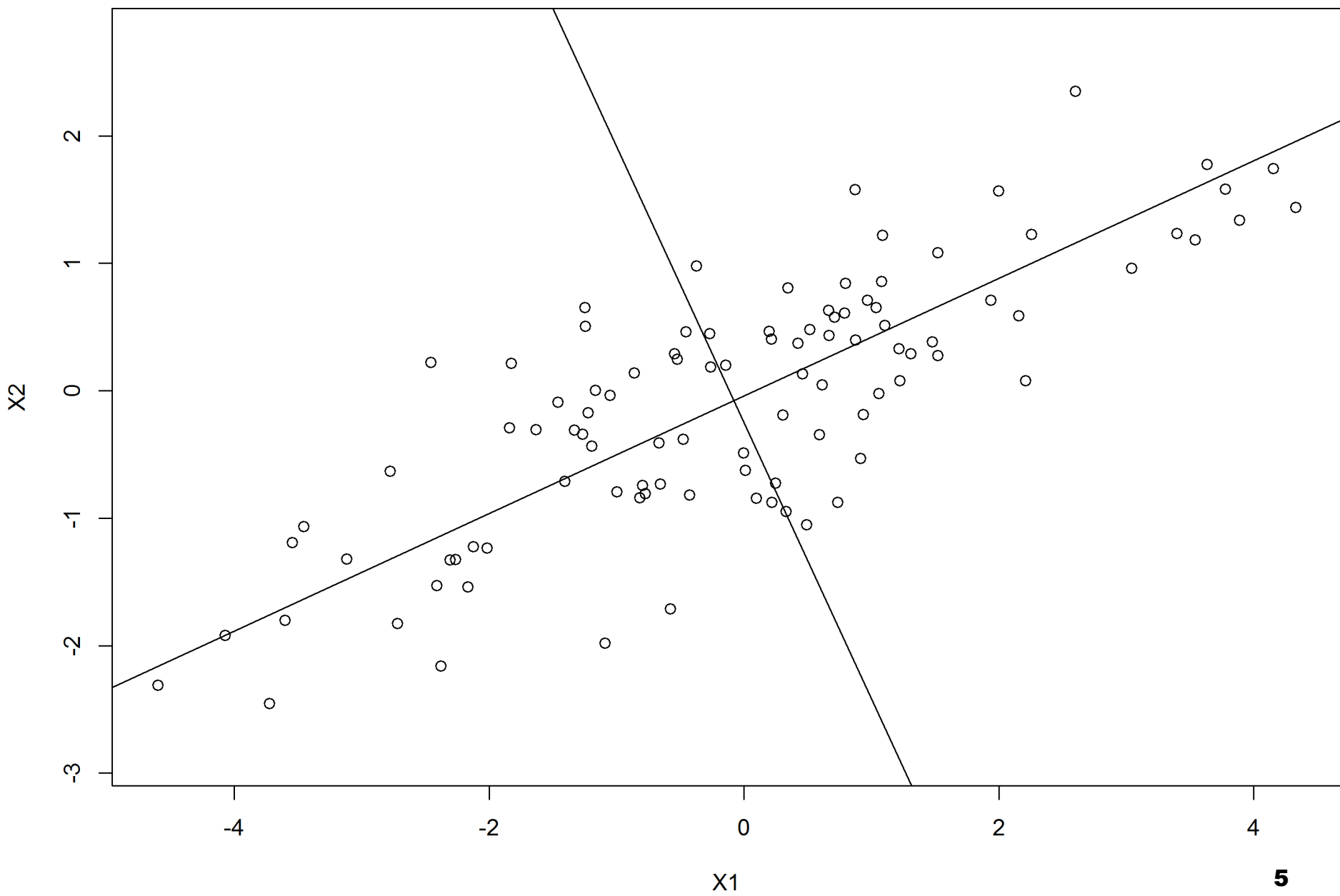
Uvod

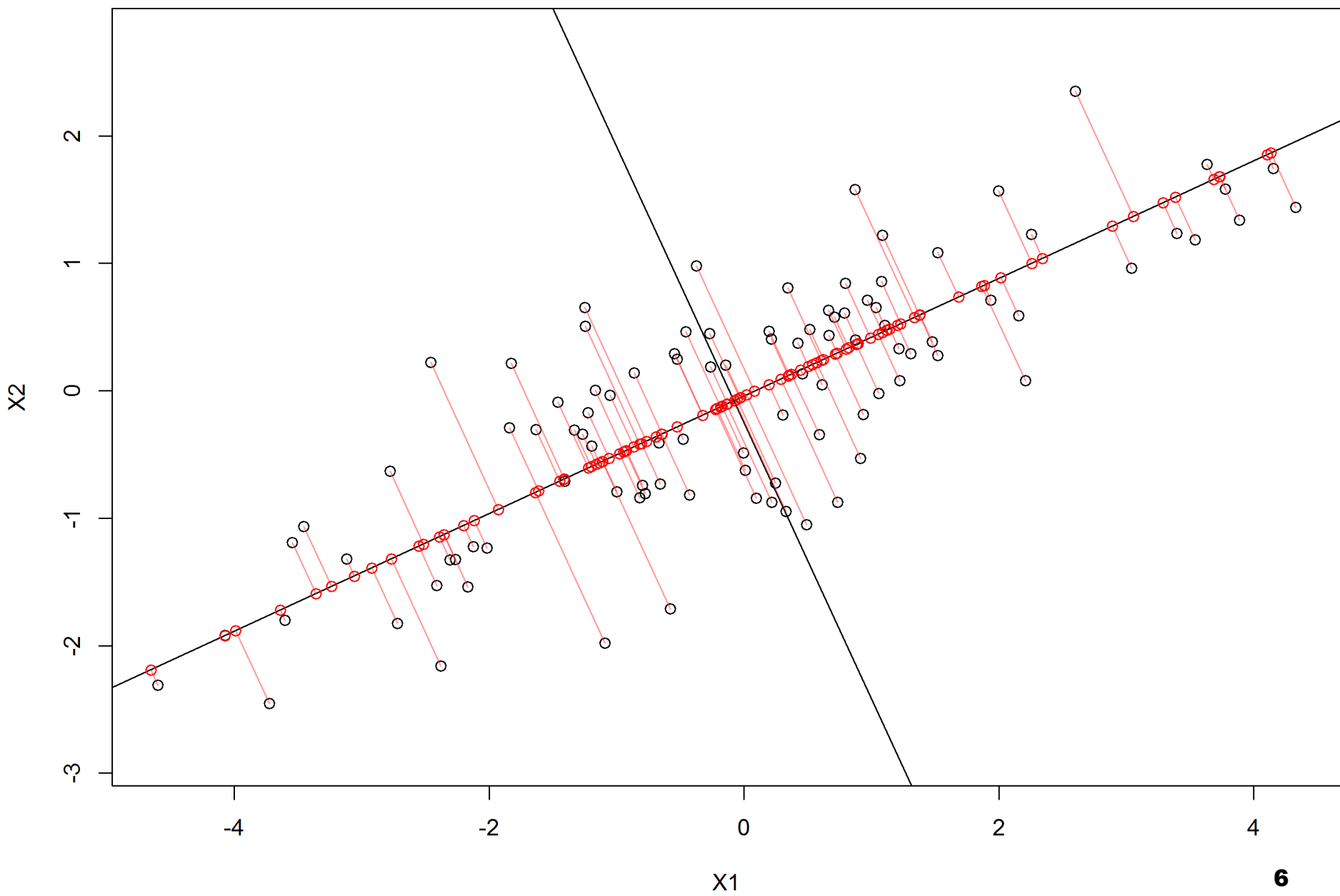
- Cilj metode je z nekaj prvimi (nekoreliranimi) glavnimi komponentami **pojasniti kar največ variabilnosti** originalnih spremenljivk
- Glavne komponente so linearne kombinacije merjenih (originalnih) spremenljivk
- Glavne komponente so med seboj neodvisne (nekorelirane)
- Komponente so urejene od najpomembnejše do najmanj pomembne (večja pomembnost pomeni, več pojasnjene variabilnosti osnovnih podatkov)

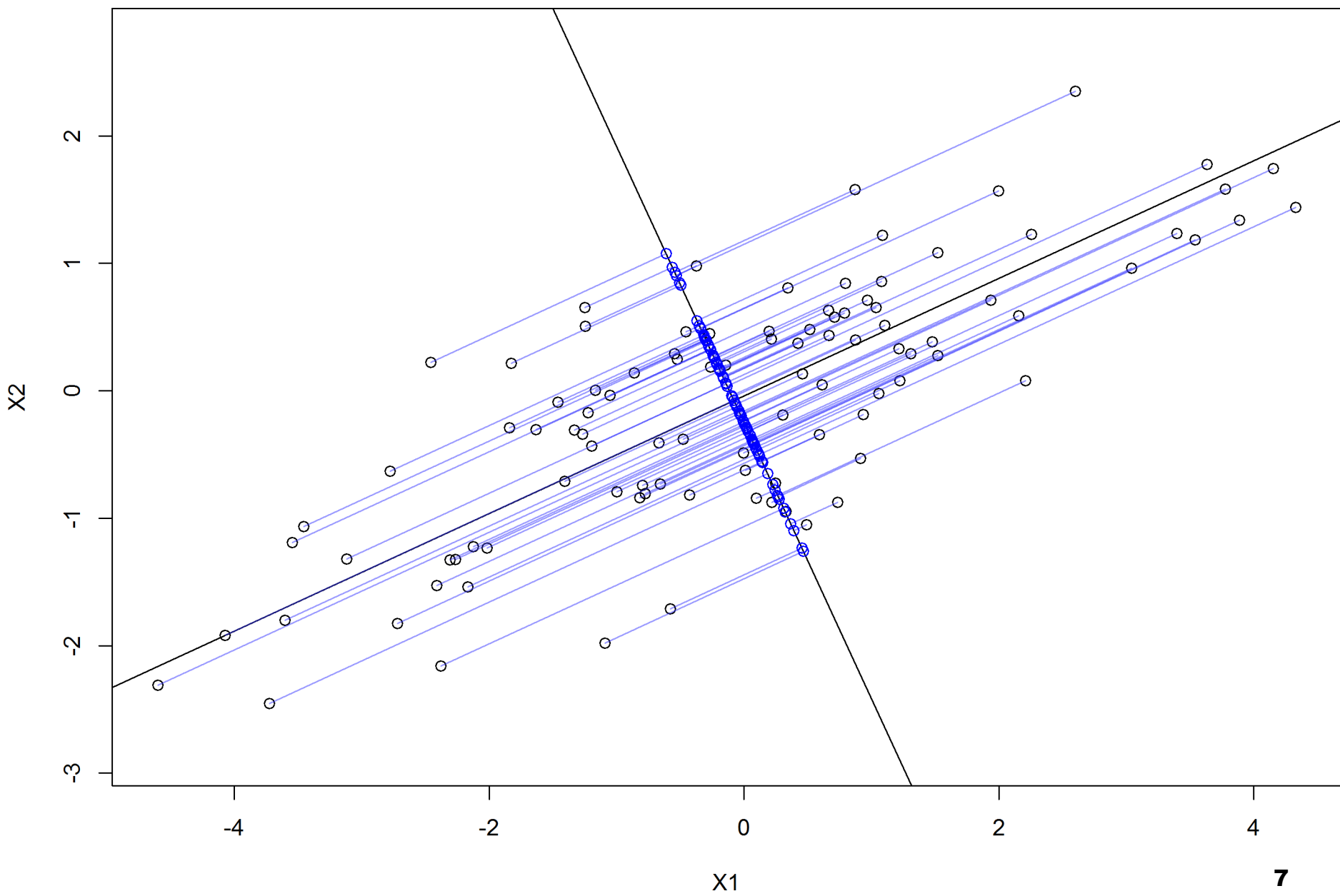
Cilj

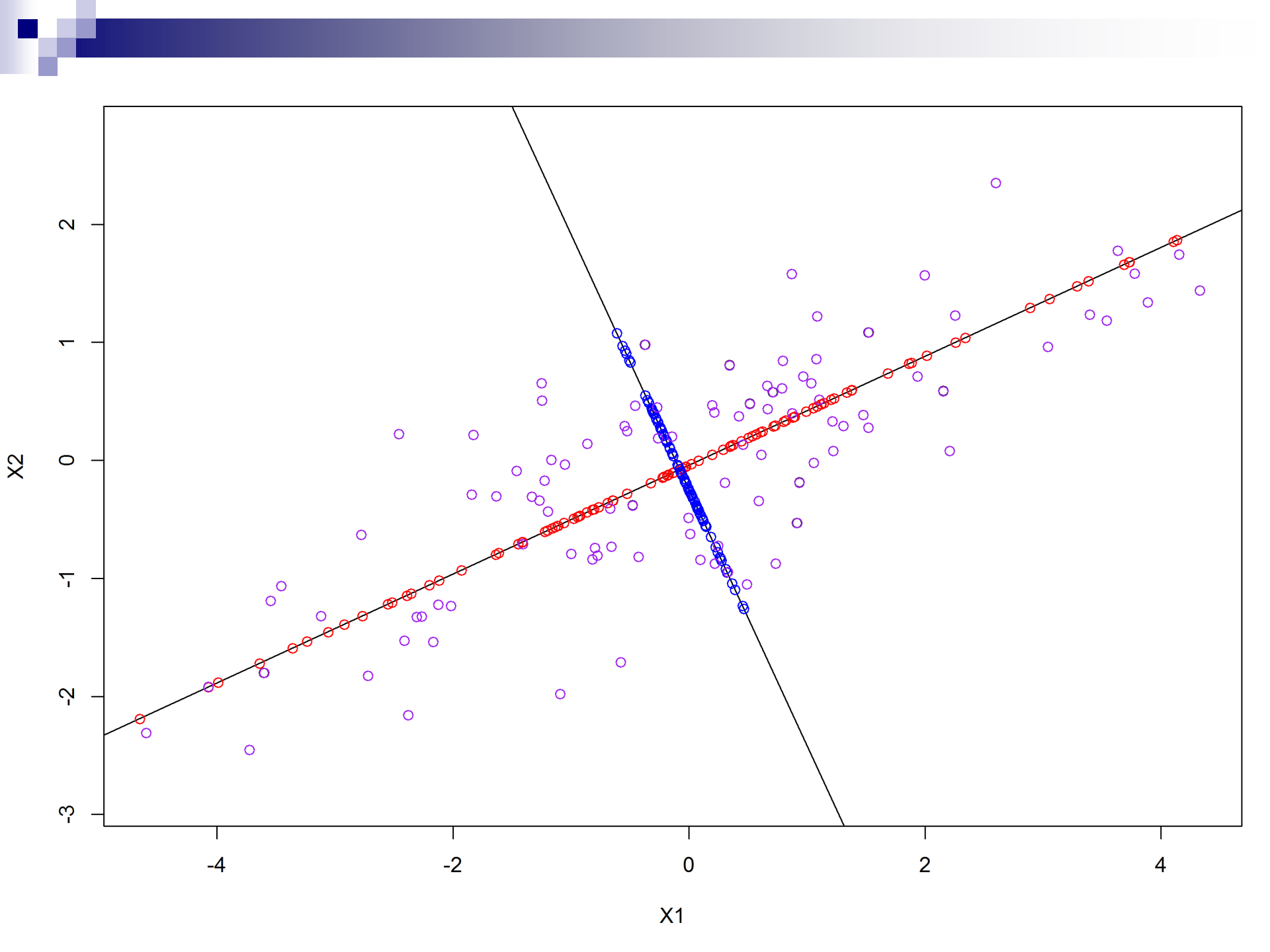
- Običajni cilj metode glavnih komponent je poiskati nekaj prvih komponent, ki pojasnjujejo večji del variabilnosti analiziranih podatkov.
- Analiza glavnih komponent omogoča povzeti podatke s čim manjšo izgubo informacij, tako da zmanjša razsežnost podatkov.
- Primer: Denimo, da želimo poiskati indeks ekonomske razvitosti občin ali držav, če razvitost merimo z več spremenljivkami ali indikatorji ekonomske razvitosti.

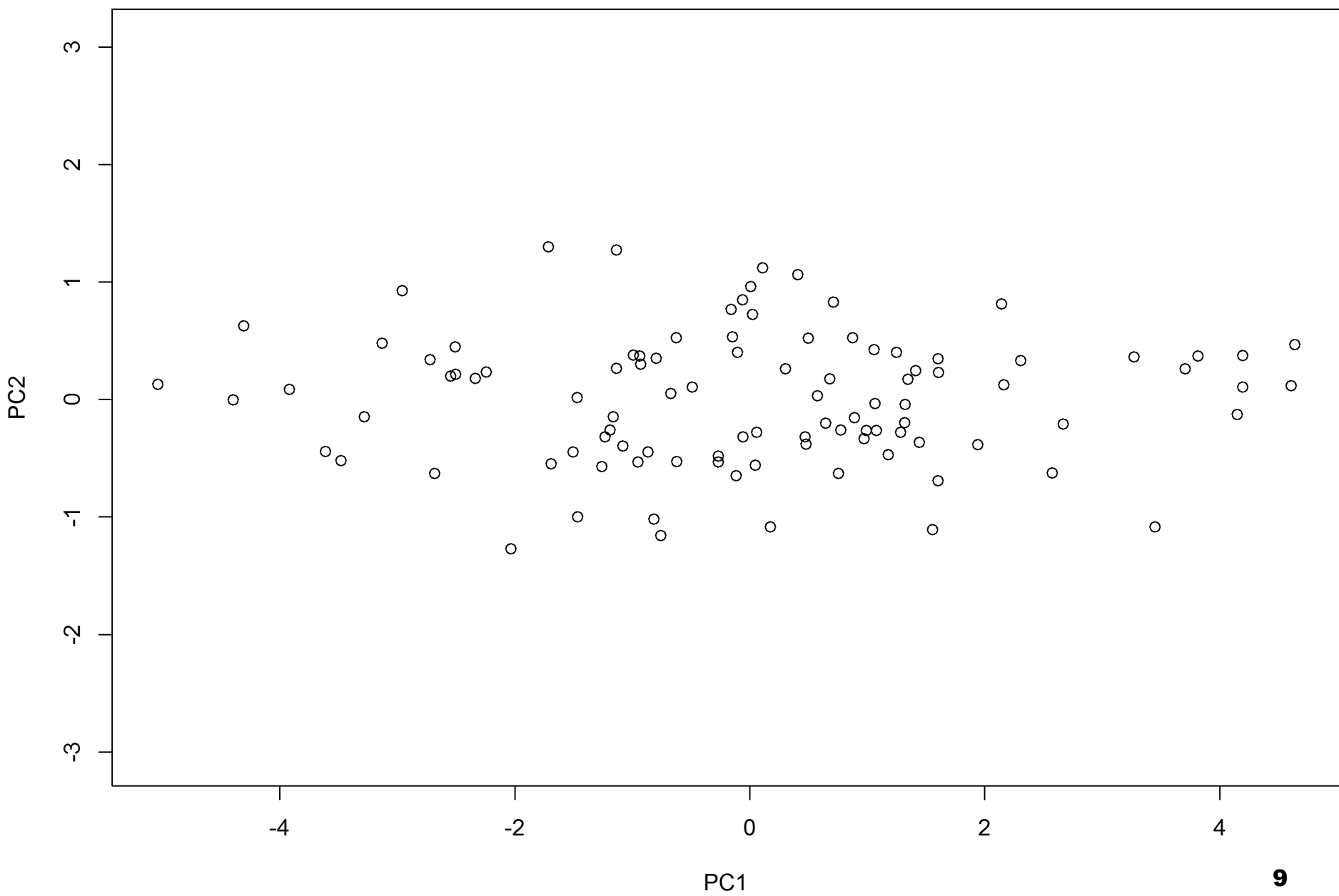












Formulacija problema

- Osnovna misel metode glavnih komponent je, da želimo poiskati take linearne kombinacije m opazovanih spremenljivk, da pojasnijo kar se da veliko variabilnosti opazovanih spremenljivk.
- Zato pri metodi glavnih komponent določimo uteži pri linearni kombinaciji spremenljivk, tako da je varianca te linearne kombinacije največja.
- Če linearno kombinacijo Y_1 opazovanih spremenljivk X_i zapišemo:

$$Y_1 = a_{11}X_1 + a_{12}X_2 + \cdots a_{1m}X_m$$

- Iščemo take uteži, da bo:

$$\text{var}(Y_1) = \max$$

Matrično

$$Y_1 = Xa_1$$

Matrika podatkov $X_{n \times m}$:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

Vektor uteži:

$$a_1 = \begin{bmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1m} \end{bmatrix}$$

- Želimo poiskati take uteži a_1 , za katere bo varianca Y_1 največja:

$$\text{var}(Y_1) = \text{var}(Xa_1) = \max$$

- Linearna kombinacija Y_1 je prva komponenta. Zato, da je problem enolično rešljiv, je potrebno podati še dodatni pogoj:

$$\sum_{i=1}^m a_{1i}^2 = a_1' a_1 = 1$$

- Ko izračunamo prvo komponento z največjo varianco, poiščemo drugo komponento tako, da je nekorelirana s prvo in ima zopet največjo varianco, oz.

$$Y_2 = a_{21}X_1 + a_{22}X_2 + \cdots a_{2m}X_m$$

$$a_2' a_2 = 1$$

$$a_2' a_1 = 0$$

- Postopek na ta način nadaljujemo in j -ta komponenta je tedaj

$$Y_j = X a_j$$

s pogoji

$$a_j' a_j = 1$$

$$a_j' a_i = 0, \text{ za vsak } i < j$$

Izračun uteži A

- Uteži za vse komponente lahko združimo v matriko A
$$A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$$
- Problem, ki ga je potrebno rešiti, je:
$$\text{var}(Y_1) = \text{var}(Xa_1) = a_1' \Sigma a_1 = \max$$

pri pogoju $a_j' a_j = 1$,
kjer je Σ matrika kovarianc (oz. korelacij, če so spremenljivke standardizirane)
opazovanih spremenljivk.
- Gre za maksimizacijo z omejitvijo, ki se ponavadi rešuje z metodo Lagrangevega multiplikatorja.

- S pomočjo odvajanja pridemo do sledeče enačbe:

$$(\Sigma - \lambda I)A = 0$$

- Ker je A različen od 0, mora biti matrika $(\Sigma - \lambda I)$ singularna. To pomeni, da:

$$q(\lambda) = |\Sigma - \lambda I| = 0$$

- $q(\lambda)$ je karakteristična enačba, njene rešitve pa so lastne vrednosti λ_i matrike Σ in a_i pripadajoči lastni vektorji.
- Lastni vektor a_1 , ki pripada največji lastni vrednosti λ_1 , nam daje uteži za iskano prvo komponento, ki je tedaj $Y_1 = Xa_1$.
- Lastni vektorji so pravokotni med seboj. Naslednji lastni vektor, ki pripada naslednji največji lastni vrednosti, podaja uteži druge komponente ...

- Kovariančno matriko Σ torej razcepimo lastne vrednosti in lastne vektorje.

$$\Sigma = A\Lambda A^T$$

- Kjer je A matrika lastnih vektorjev, lastni vektorji predstavljajo stolpce te matrike.
- Λ je diagonalna matrika, ki ima na diagonalni lastne vrednosti (λ_i) , druge pa 0-le.

Singularni razcep (SVD)

- Enakovredno rešitev se lahko dobi, če razcepimo matriko podatkov X z razcepom SVD

$$X = USV^T$$

- $S^2 / (n - 1) = \Lambda$ je tu diagonalna matrika lastnih vrednosti
- $V = A$ je matrika (nereskaliranih uteži)

Značilnosti rešitve

- Lastne vrednosti λ_i so variacije glavnih komponent
- Vsota variance merjenih spremenljivk je enaka vsoti lastnih vrednosti (oz. varianc glavnih komponent):

$$\sum_{i=1}^m \sigma_{ii}^2 = \sum_{i=1}^m \lambda_i$$

Variabilnost se torej ohranja.

- Dobljeni rezultati metode glavnih komponent so smiselni, če so variance glavnih komponent λ_i pozitivna števila.
- Ta pa so, če je matrika Σ pozitivno definitna. To pa je, če je rang matrike podatkov X enak številu merjenih spremenljivk m .
- To pomeni, da nobena spremenljivka ne sme biti linearna kombinacija ostalih spremenljivk.
- Tudi vsaj m vrstic mora biti linearno neodvisnih. To pomeni, da moramo imeti za vsak slučaj precej več enot kot spremenljivk.

- Delež skupne variance, ki jo pojasni j -ta glavna komponenta je potem:

$$\frac{\lambda_j}{\sum_{i=1}^m \sigma_{ii}^2} = \frac{\lambda_j}{\sum_{i=1}^m \lambda_i}$$

- Če lahko merjene spremenljivke v različnih merskih enotah ali lestvicah, je njihova linearna kombinacija nesmiselna.
- Zato, predno računamo glavne komponente, spremenljivke standardiziramo.
- To pomeni, da je matrika varianc in kovarianc korelacijska matrika.
- Ker so v diagonali matrike enice, je delež pojasnjene variance z j -to glavno komponento enak: λ_j / m

Še nekaj lastnosti - matrično

Matrika podatkov $X_{n \times m}$
(naj bo centrirana):

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1m} \\ x_{21} & x_{22} & \cdots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nm} \end{bmatrix}$$

Matrika uteži $A_{m \times m}$
(nereskalirana):

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mm} \end{bmatrix}$$

Matrika glavnih
komponent $Y_{n \times m}$:

$$Y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1m} \\ y_{21} & y_{22} & \cdots & y_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nm} \end{bmatrix}$$

Potem velja:

$$Y = XA$$

in tudi

$$X = YA'$$

Omejitev na samo nekaj komponent

Matriki komponent (Y) in utež (A) lahko tudi omejimo le na nekaj prvih glavnih komponent.

Na primer, če je A_L matrika uteži le za prvih L komponent (dimenzije $m \times L$) in matrika Y_L matrika prvih L komponent (dimenzije $n \times L$), potem velja:

$$Y_L = X A_L$$

in tudi

$$X^* = Y_L A_L'$$

kjer je X^* približek matrike X na osnovi prvih L komponent.

Reskaliranje uteži

- Ponavadi dobljene glavne komponente reskaliramo tako, da je posamezna reskalirana utež: $a_{ij}^* = a_{ij}\sqrt{\lambda_i}$
- Med tem ko velja $\sum_{j=1}^m a_{ij}^2 = a_i' a_i = 1$, velja $\sum_{j=1}^m (a_{ij}^*)^2 = \lambda_i \rightarrow$ vsota reskaliranih uteži po komponentah so lastne vrednosti
- Utež reskalirane i -te glavne komponente a_{ij}^* je **korelacija** (če smo analizirali korelacijsko matriko) med i -to komponento in j -to spremenljivko. \rightarrow S pomočjo teh korelacij interpretiramo glavne komponente
- Velja tudi: $\sum_{i=1}^m (a_{ij}^*)^2 = \sigma_{jj} \rightarrow$ Vsota reskaliranih uteži po spremenljivkah je varianca spremenljivke.

Kdaj uporabiti metodo glavnih komponent

- Metodo glavnih komponent je smiselno uporabiti, če spremenljivke korelirajo med seboj.
- Če ne korelirajo, so komponente kar merjene spremenljivke.
- Zato je pred uporabo metode glavnih komponent zaželeno zavrniti naslednjo ničelno domnevo:

$$H_0: \Sigma = I$$

kar pomeni, da so vse opazovane spremenljivke nekorelirane med seboj.

- Bartlett je, ob predpostavki večrazsežne normalne porazdelitve merjenih spremenljivk, našel naslednjo statistiko za preverjanje zgornje ničelne domneve:

$$\chi^2 = - \left(n - 1 - \frac{2m + 5}{6} \right) \ln |R|$$

- ki se asimptotično porazdeljuje po χ^2 porazdelitvi z $\frac{m(m-1)}{2}$ prostostnimi stopinjami, kjer je R vzorčna korelacijska matrika.
- Z Bartlettovim testom, ki je bil še nadalje razdelan, je mogoče preveriti, da je $m - k$ preostalih komponent zanemarljivih: da so njihove lastne vrednosti enake med seboj.

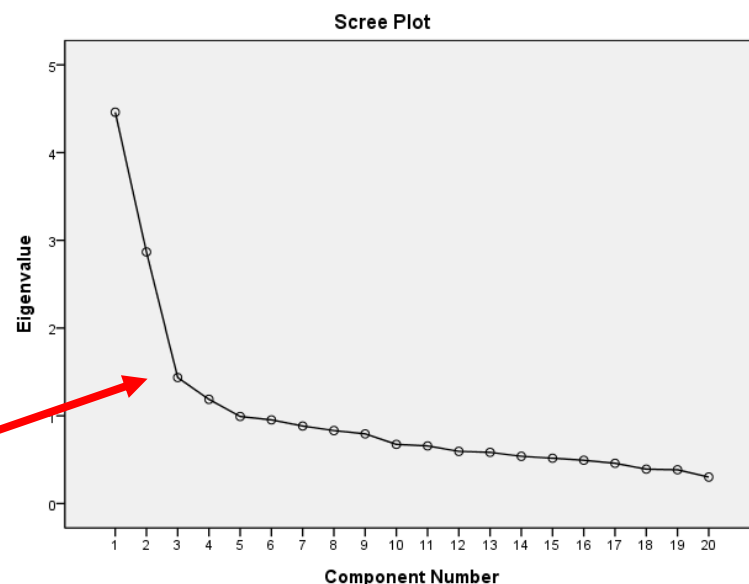
Odločitev o številu komponent

- Tako da pojasnimo vsaj ... % celotne variabilnosti (npr. vsaj 80%, vsaj 50%)
- Tako da zadnja komponenta pojasni vsaj 5% celotne variabilnosti
- Dokler so lastne vrednosti večje od povprečne variabilnosti spremenljivk (v primeru korelacijske matrike oz. standardiziranih spremenljivke torej večje od 1)
- **Tam kjer je “koleno” scree diagramu (komponente, kjer je koleno ne izberemo več).**
- **Primerjava s lastnimi vrednostmi na podlagi podatkov, slučajno generiranih iz neodvisnih (ne-koreliranih spremenljivk).**

„Scree“ diagram (diagram lastnih vrednosti)

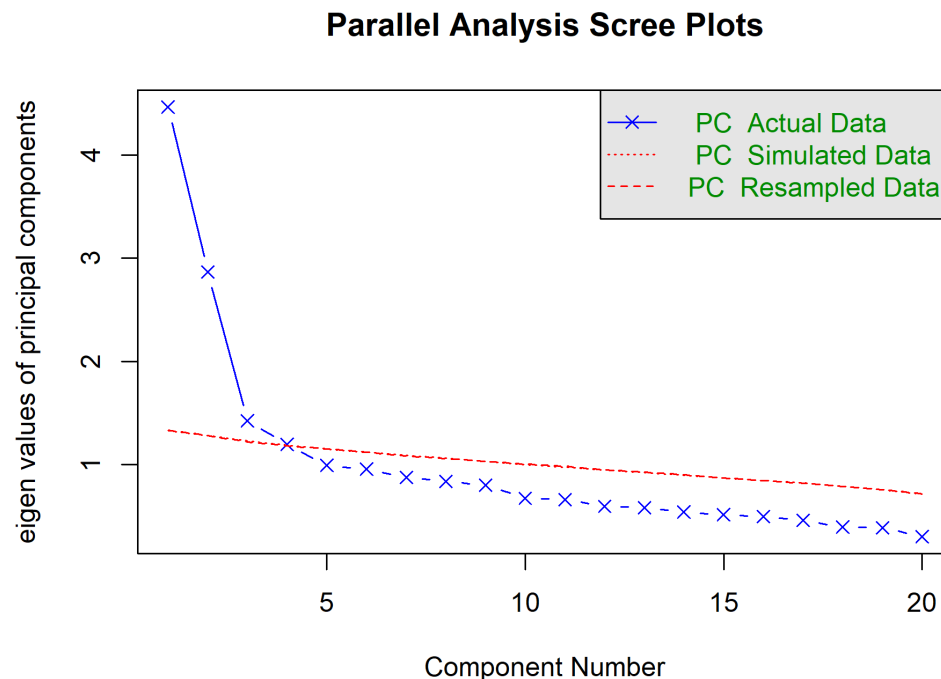
- Grafične predstavitve lastnih vrednosti.
- Na abscisno os nanašamo zaporedne številke komponent, na ordinatno pa ustrezne lastne vrednosti.
- Tam kjer se graf „lomi“ oz. je koleno je priporočilo za število komponent. Obdržimo komponente do kolena (tiste pri kolenu ne upoštevamo več).

koleno



Paralelna analiza

- Primerjava empiričnih lastnih vrednosti s pričakovani
- Pričakovane vrednosti izračunamo s simulacijami na nekoreliranih spremenljivkah
- Izberemo komponente, kjer je empirična lastna vrednost večja od pričakovane



Povzetek

- Glavni rezultat so uteži glavnih komponent
- Če jih pravilno “reskaliramo” dobimo korelacije med glavnimi komponentami in spremenljivkami (SPSS vrne reskalirane uteži, v R-ju `princomp` in `prcomp` ne, `principal` {psych} pa ja)
- Če seštejemo kvadrate “reskaliranih” uteži po komponentah, dobimo lastne vrednosti komponent, po spremenljivkah pa njihove variance (če za vse komponente) oz. variabilnost, pojasnjeno s prvimi k komponentami (komunalitete), če le za prvih k komponent.
- Ker so glavne komponente linearne kombinacije merjenih spremenljivk, jih je enostavno izračunati.

Primer

- Podatki so bili zbrani v okviru raziskave *Kakovost merjenja egocentričnih socialnih omrežij* (Ferligoj in drugi, 2000) leta 2000. Vzorec vsebuje 1033 prebivalcev Ljubljane. Analiza je bila narejena na 631 prebivalcih, ki so bili osebno intervjujani.
- Za analizo glavnih komponent smo izbrali spremenljivke, ki merijo Ekstravertiranost in Emocionalno stabilnost (morda bi bilo bolje sicer vzeti samo en sklop, a smo vzeli oba zaradi navezave na primer z razvrščanja).

Korelacije

	EKSTA	EKSTB	EKSTE	EKSTH	EKSTIR	EKSTLR	EKSTNR	EKSTOR	EKSTP	EKSTRR	EMOCC	EMOCDR	EMOCF	EMOCGR	EMOCJR	EMOCKR	EMOCMR	EMOCQR	EMOCSR	EMOCTR
EKSTA Sem duša vsake družbe.		,203	,282	,220	,127	,286	,195	,160	,165	,270	,055	,056	,262	-,026	-,037	-,033	-,010	-,003	,035	,005
EKSTB Ne moti me, če sem v središču pozornosti.	,203		,186	,182	,237	,150	,192	,152	,158	,289	-,015	,076	,141	,003	,008	,100	-,001	-,055	,007	,019
EKSTE Na zabavah se pomenkujem z mnogo ljudmi vseh vrst	,282	,186		,267	,044	,313	,334	,235	,304	,285	,052	,102	,283	,007	,049	,106	,002	,055	,143	,084
EKSTH Pogovore nacenjam jaz.	,220	,182	,267		,038	,284	,279	,237	,244	,212	,091	,037	,135	-,023	-,065	,090	-,022	-,003	,049	-,035
EKSTIR Nerač/a pritegnem pozornost nase.	,127	,237	,044	,038		,155	,165	,165	-,009	,324	-,070	,105	,040	,101	,094	,130	-,088	-,054	,047	,062
EKSTLR Sem redkobeseden(a).	,286	,150	,313	,284	,155		,365	,469	,237	,408	,060	-,013	,287	-,036	-,026	,134	,046	,041	,147	,063
EKSTNR V navzočnosti neznanih oseb sem molcec(a).	,195	,192	,334	,279	,165	,365		,425	,200	,459	,061	,131	,258	,143	,126	,250	,039	,108	,160	,195
EKSTOR Imam malo povedati.	,160	,152	,235	,237	,165	,469	,425		,234	,444	,035	,084	,225	,025	,067	,174	,054	,121	,143	,139
EKSTP Med ljudmi se pocutim sproščeno.	,165	,158	,304	,244	-,009	,237	,200	,234		,193	,142	,117	,482	,080	,095	,085	,049	,060	,147	,067
EKSTRR Zadržujem se v ozadju.	,270	,289	,285	,212	,324	,408	,459	,444	,193		,089	,142	,185	,077	,145	,276	,072	,164	,272	,215
EMOCC Redkokdaj sem potr(a).	,055	-,015	,052	,091	-,070	,060	,061	,035	,142	,089		,223	,215	,173	,162	,205	,158	,221	,351	,171
EMOCDR Zlahka me kaj vrže iz tira.	,056	,076	,102	,037	,105	-,013	,131	,084	,117	,142	,223		,203	,588	,588	,294	,223	,369	,337	,431
EMOCF Sem vecidel sproščen(a).	,262	,141	,283	,135	,040	,287	,258	,225	,482	,185	,215	,203		,148	,184	,230	,085	,194	,266	,214
EMOCGR Zlahka me kaj razdraži.	-,026	,003	,007	-,023	,101	-,036	,143	,025	,080	,077	,173	,588	,148		,647	,255	,271	,359	,254	,365
EMOCJR Zlahka me kaj vznemiri.	-,037	,008	,049	-,065	,094	-,026	,126	,067	,095	,145	,162	,588	,184	,647		,305	,225	,336	,376	,448
EMOCKR Sem zaskrbljene narave.	-,033	,100	,106	,090	,130	,134	,250	,174	,085	,276	,205	,294	,230	,255	,305		,088	,250	,350	,391
EMOCMR Velikokrat sem muhasto razpoložen(a).	-,010	-,001	,002	-,022	-,088	,046	,039	,054	,049	,072	,158	,223	,085	,271	,225	,088		,409	,271	,267
EMOCQR Moje razpoloženje se pogosto menja.	-,003	-,055	,055	-,003	-,054	,041	,108	,121	,060	,164	,221	,369	,194	,359	,336	,250	,409		,439	,442
EMOCSR Pogosto sem potr(a).	,035	,007	,143	,049	,047	,147	,160	,143	,147	,272	,351	,337	,266	,254	,376	,350	,271	,439		,379
EMOCTR Zlahka se me poloti napetost.	,005	,019	,084	-,035	,062	,063	,195	,139	,067	,215	,171	,431	,214	,365	,448	,391	,267	,442	,379	

Barlettov test

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,833
Bartlett's Test of Sphericity	Approx. Chi-Square	3153,239
	df	190
	Sig.	,000

- Ker je $p < 0,05$, ničelno domnevo zavrnamo → sklep: vsaj nekatere spremenljivke so korelirane

Lastne vrednosti in % pojas. var.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	4,460	22,301	22,301	4,460	22,301	22,301
2	2,867	14,337	36,638	2,867	14,337	36,638
3	1,436	7,180	43,818	1,436	7,180	43,818
4	1,188	5,939	49,757	1,188	5,939	49,757
5	,992	4,960	54,717	,992	4,960	54,717
6	,954	4,769	59,486	,954	4,769	59,486
7	,884	4,418	63,904	,884	4,418	63,904
8	,831	4,157	68,061	,831	4,157	68,061
9	,794	3,970	72,031	,794	3,970	72,031
10	,674	3,371	75,402	,674	3,371	75,402
11	,656	3,282	78,684	,656	3,282	78,684
12	,595	2,974	81,658	,595	2,974	81,658
13	,584	2,919	84,577	,584	2,919	84,577
14	,538	2,692	87,269	,538	2,692	87,269
15	,516	2,582	89,851	,516	2,582	89,851
16	,493	2,467	92,318	,493	2,467	92,318
17	,458	2,291	94,609	,458	2,291	94,609
18	,392	1,958	96,567	,392	1,958	96,567
19	,385	1,924	98,491	,385	1,924	98,491
20	,302	1,509	100,000	,302	1,509	100,000

Extraction Method: Principal Component Analysis.

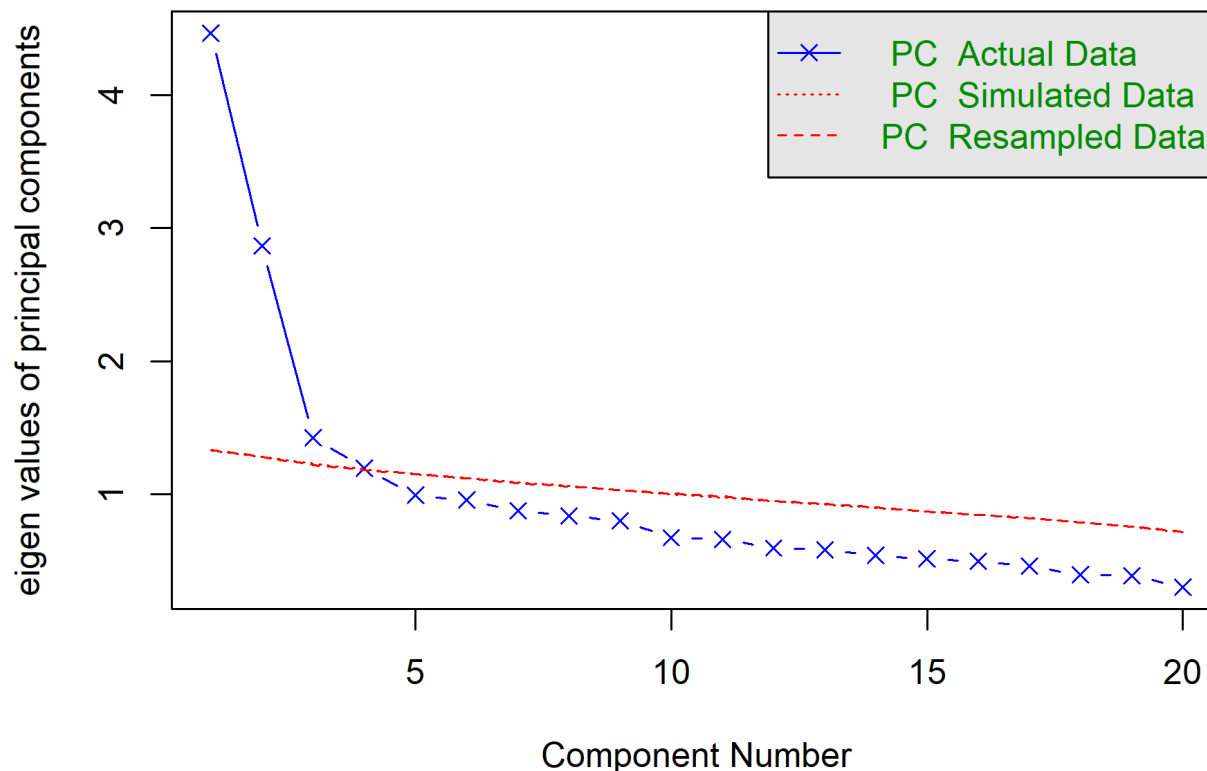
Scree plot



Koleno je pri **3** komponentah, zato obdržimo **2** komponenti

Paralelna analiza

Parallel Analysis Scree Plots



- Tukaj predlaga 4 komponente (čeprav ni dobro vidno, ker je 4 ravno na meji)

Uteži (reskalirane)

Component Matrix^a

	Component																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
EKSTA Sem duša vsake družbe.	,274	,438	-,089	,253	,324	,360	-,073	-,466	-,177	,305	-,070	-,034	-,031	-,129	,126	,102	,082	,049
EKSTB Ne moti me, če sem v središču pozornosti.	,252	,360	,282	,305	,075	,412	,056	,560	-,081	,008	-,237	-,158	,130	,002	-,163	-,070	-,019	,044
EKSTH Na zabavah se pomenkujem z mnogo ljudmi vseh vrst	,409	,423	-,190	,163	,098	-,123	,078	-,067	-,480	-,468	,073	,118	,131	,220	-,010	,120	,025	,077
EKSTI Pogovore nacenjam jaz.	,277	,449	-,162	,063	,024	-,030	,681	,072	,099	,212	,366	-,032	-,048	,005	-,010	-,100	,049	,071
EKSTJ Nerač/a pritegnem pozornost nase.	,227	,212	,663	,117	-,096	,254	-,198	-,104	,241	-,108	,363	,215	-,139	,218	-,049	,003	,062	,018
EKSTK Sem redkobeseden(a).	,429	,544	-,045	-,257	,081	-,092	-,099	-,163	,183	,061	,015	,088	,453	,023	-,112	-,203	-,278	-,101
EKSTL V navzočnosti neznanih oseb sem molčec(a).	,540	,386	,156	-,146	,003	-,271	,153	-,012	-,087	-,036	-,304	,248	-,388	-,126	-,152	-,154	,048	-,113
EKSTM Imam malo povedati.	,476	,431	,107	-,328	,063	-,281	-,068	,012	,278	,031	-,147	-,289	,067	,166	-,005	,219	,316	,120
EKSTN Med ljudmi se počutim sproščeno.	,408	,300	-,432	,357	-,095	-,186	-,213	,270	,223	-,081	,146	-,074	-,145	-,047	,292	,034	-,124	-,058
EKSTO Zadržujem se v ozadju.	,580	,390	,302	-,236	,010	,133	-,042	-,018	-,002	-,146	-,048	-,131	-,141	-,198	,267	,071	-,235	-,027
EMOCC Redkokdaj sem potr(a).	,364	-,176	-,401	-,053	-,371	,428	,229	-,190	,272	-,141	-,309	,060	-,053	,240	,044	,005	-,056	,063
EMOCDR Zlahka me kaj vrže iz tira.	,601	-,426	,131	,323	,079	-,051	,148	-,101	,034	-,001	-,025	-,149	,083	,110	-,010	,089	,085	-,478
EMOCF Sem vecidel sproščen(a).	,542	,187	-,386	,277	-,160	-,044	-,403	,036	,026	,213	,003	,136	-,061	,000	-,263	-,035	,099	,030
EMOCGR Zlahka me kaj razdraži.	,529	-,505	,180	,309	,159	-,181	,136	-,087	,190	-,008	-,080	,127	,029	-,071	-,075	,057	-,153	,163
EMOCJR Zlahka me kaj vznemiri.	,578	-,501	,191	,261	,034	-,184	,016	-,093	,062	-,097	-,044	-,066	,127	-,168	,054	-,097	,035	,257
EMOCKR Sem zaskrbljene narave.	,543	-,134	,164	-,154	-,509	-,018	,070	,155	-,234	,274	,035	,238	,193	-,104	,104	,307	,003	-,004
EMOCMR Velikokrat sem muhasto razpoložen(a).	,347	-,336	-,213	-,279	,545	,191	-,025	,288	,140	-,026	,008	,374	,064	-,010	,173	,022	,162	-,033
EMOCQR Moje razpoloženje se pogosto menja.	,534	-,409	-,167	-,291	,231	,085	-,045	,056	-,100	,070	,176	-,181	-,215	,094	-,302	,228	-,268	,046
EMOCSR Pogosto sem potr(a).	,606	-,249	-,166	-,238	-,191	,245	-,060	-,063	-,056	-,250	,223	-,163	,054	-,328	-,116	-,182	,226	-,040
EMOCTR Zlahka se me poloti napetost.	,594	-,366	,095	-,135	-,025	-,056	-,119	,039	-,296	,233	,006	-,110	-,068	,332	,233	-,365	-,001	,057

Extraction Method: Principal Component Analysis.

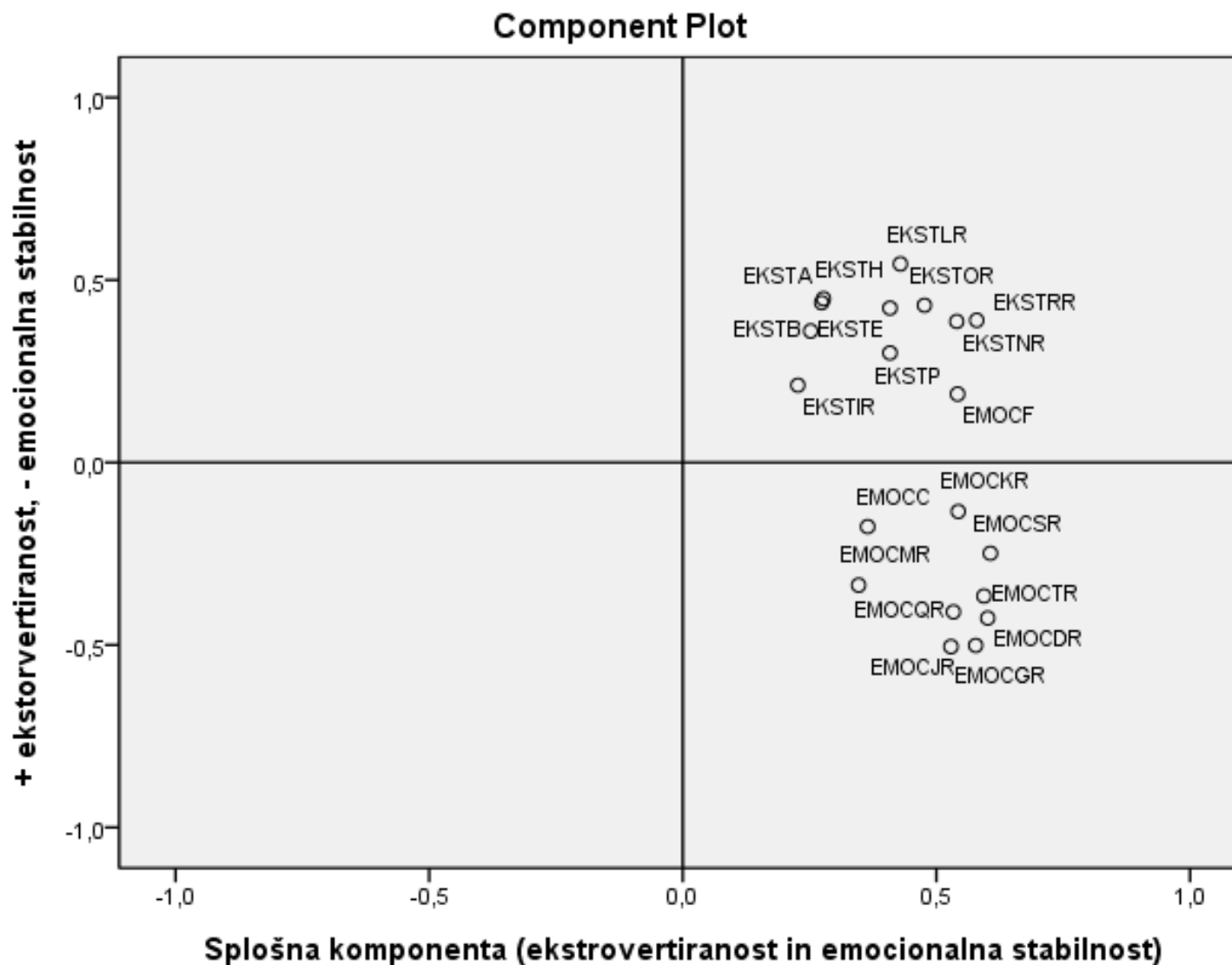
a. 20 components extracted.

Uteži (reskalirane)

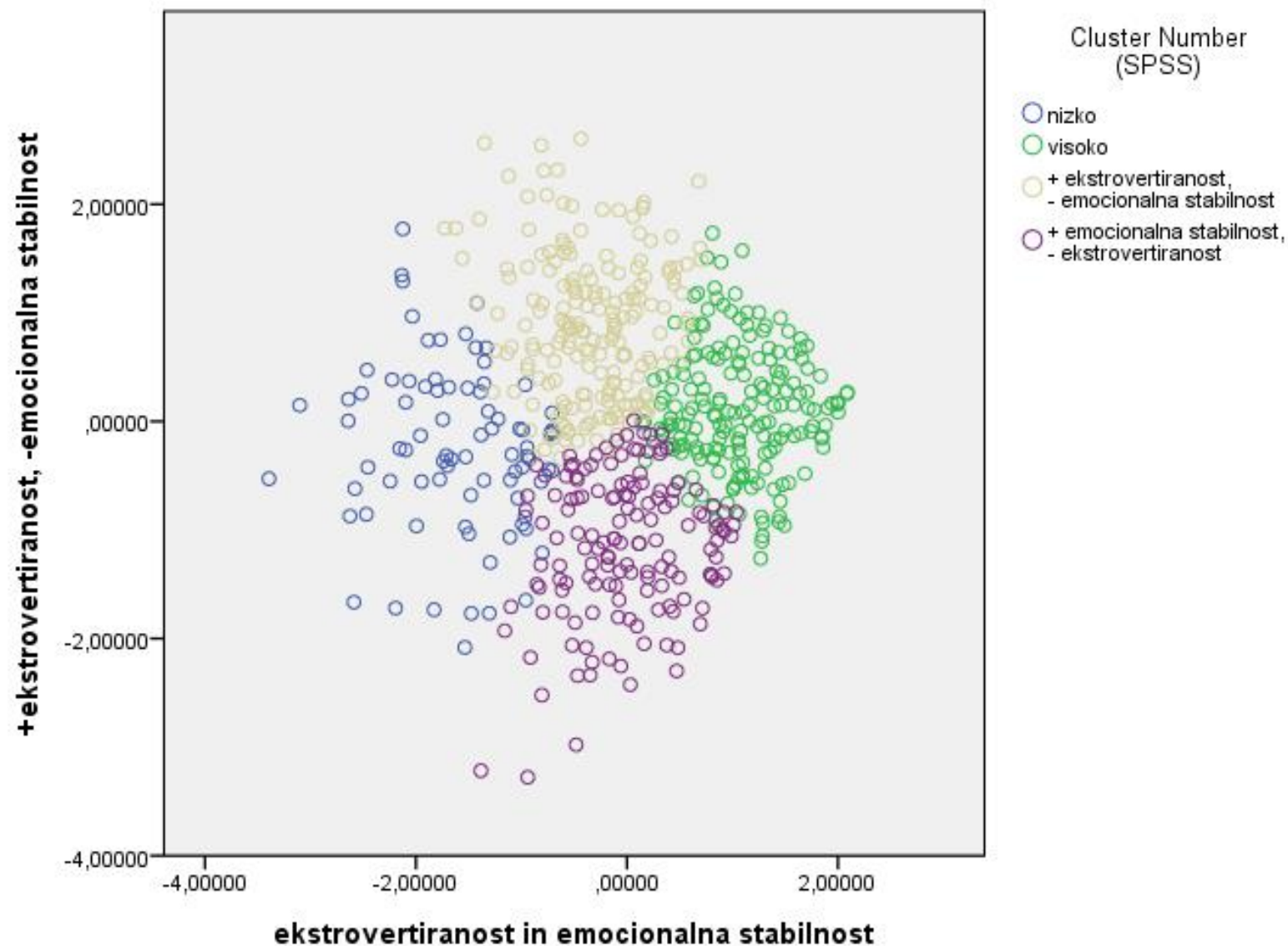
Component Matrix^a

	Component															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
EKSTA Sem duša vsake družbe.	,274	,438	-,089	,253	,324	,360	-,073	-,466	-,177	,305	-,070	-,034	-,031	-,129	,126	,102
EKSTB Ne moti me, če sem v središču pozornosti.	,252	,360	,282	,305	,075	,412	,056	,560	-,081	,008	-,237	-,158	,130	,002	-,163	-,070
EKSTE Na zabavah se pomenkujem z mnogo ljudmi vseh vrst	,409	,423	-,190	,163	,098	-,123	,078	-,067	-,480	-,468	,073	,118	,131	,220	-,010	,120
EKSTH Pogovore nacenjam jaz.	,277	,449	-,162	,063	,024	-,030	,681	,072	,099	,212	,366	-,032	-,048	,005	-,010	-,100
EKSTIR Nerad/a pritegnem pozornost nase.	,227	,212	,663	,117	-,096	,254	-,198	-,104	,241	-,108	,363	,215	-,139	,218	-,049	,003
EKSTLR Sem redkobeseden(a).	,429	,544	-,045	-,257	,081	-,092	-,099	-,163	,183	,061	,015	,088	,453	,023	-,112	-,203
EKSTNR V navzocnosti neznanih oseb sem molcec(a).	,540	,386	,156	-,146	,003	-,271	,153	-,012	-,087	-,036	-,304	,248	-,388	-,126	-,152	-,154
EKSTOR Imam malo povedati.	,476	,431	,107	-,328	,063	-,281	-,068	,012	,278	,031	-,147	-,289	,067	,166	-,005	,219
EKSTP Med ljudmi se pocutim sproščeno.	,408	,300	-,432	,357	-,095	-,186	-,213	,270	,223	-,081	,146	-,074	-,145	-,047	,292	,034
EKSTRR Zadržujem se v ozadju.	,580	,390	,302	-,236	,010	,133	-,042	-,018	-,002	-,146	-,048	-,131	-,141	-,198	,267	,071
EMOCC Redkokdaj sem potr(a).	,364	-,176	-,401	-,053	-,371	,428	,229	-,190	,272	-,141	-,309	,060	-,053	,240	,044	,005
EMOCDR Zlahka me kaj vrže iz tira.	,601	-,426	,131	,323	,079	-,051	,148	-,101	,034	-,001	-,025	-,149	,083	,110	-,010	,089
EMOCF Sem vecidel sproščen(a).	,542	,187	-,386	,277	-,160	-,044	-,403	,036	,026	,213	,003	,136	-,061	,000	-,263	-,035
EMOCGR Zlahka me kaj razdraži.	,529	-,505	,180	,309	,159	-,181	,136	-,087	,190	-,008	-,080	,127	,029	-,071	-,075	,057
EMOCJR Zlahka me kaj vznemiri.	,578	-,501	,191	,261	,034	-,184	,016	-,093	,062	-,097	-,044	-,066	,127	-,168	,054	-,097
EMOCKR Sem zaskrbljene narave.	,543	-,134	,164	-,154	-,509	-,018	,070	,155	-,234	,274	,035	,238	,193	-,104	,104	,307
EMOCMR Velikokrat sem muhasto razpoložen(a).	,347	-,336	-,213	-,279	,545	,191	-,025	,288	,140	-,026	,008	,374	,064	-,010	,173	,022
EMOCQR Moje razpoloženje se pogosto menja.	,534	-,409	-,167	-,291	,231	,085	-,045	,056	-,100	,070	,176	-,181	-,215	,094	-,302	,228
EMOCSR Pogosto sem potr(a).	,606	-,249	-,166	-,238	-,191	,245	-,060	-,063	-,056	-,250	,223	-,163	,054	-,328	-,116	-,182
EMOCTR Zlahka se me poloti napetost.	,594	-,366	,095	-,135	-,025	-,056	-,119	,039	-,296	,233	,006	-,110	-,068	,332	,233	-,365

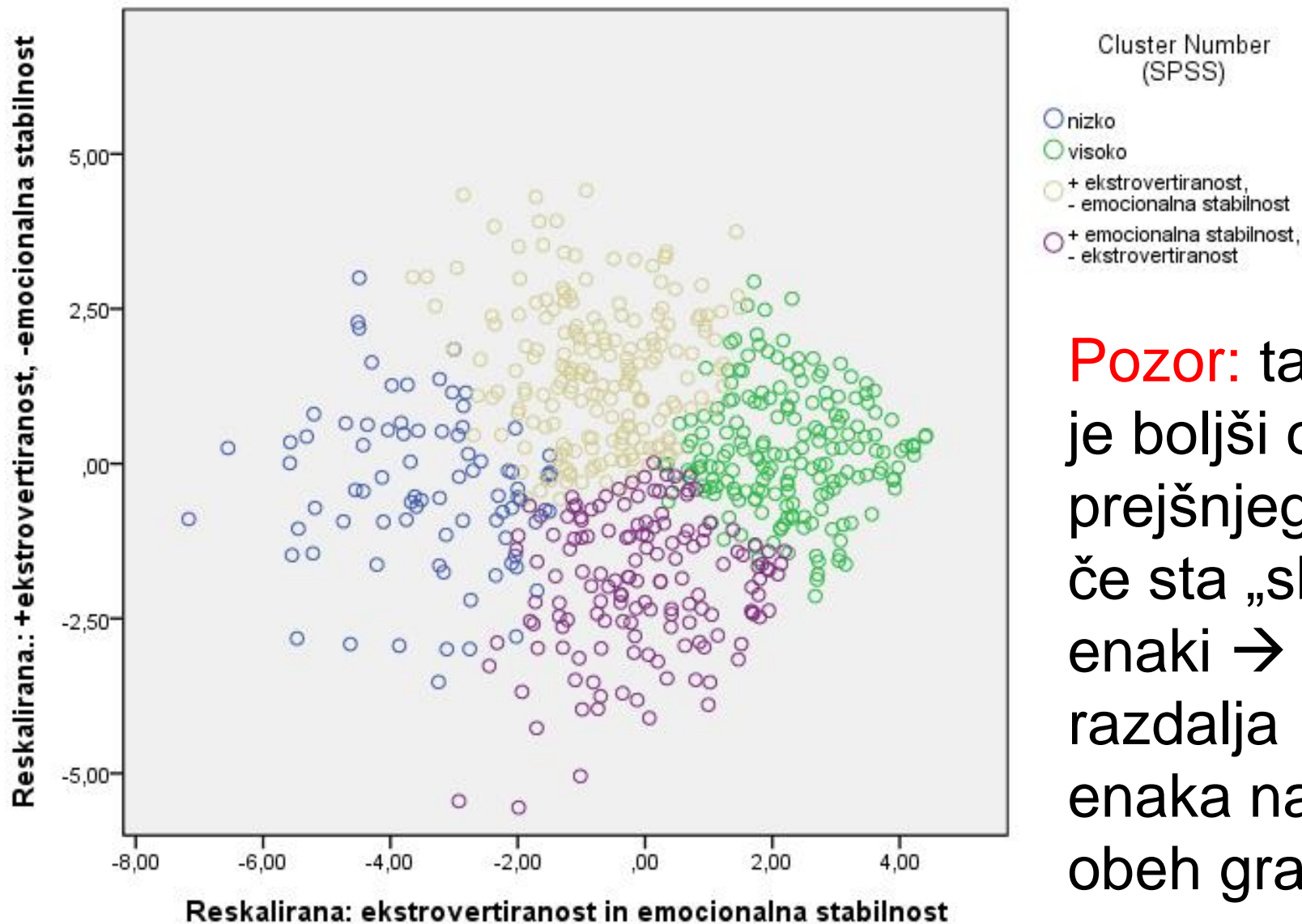
Grafična predstavitev uteži



Grafična predstavitev podatkov



GPP – reskalirane komponente

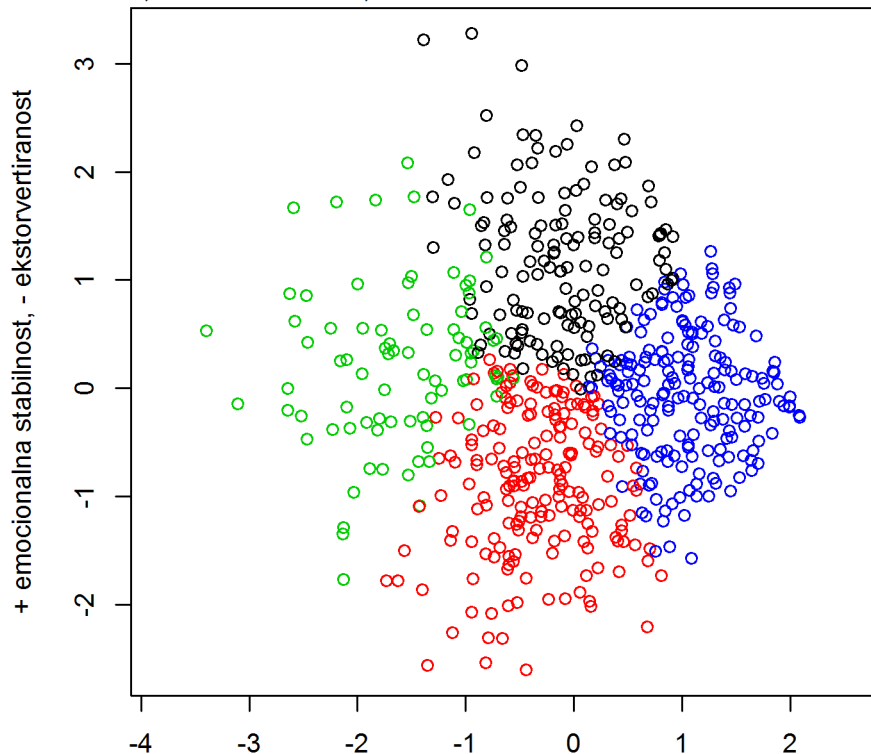


Pozor: ta graf je boljši od prejšnjega le, če sta „skali“ enaki → če je razdalja 1 enaka na obeh grafih.

Vpliv reskaliranja na komponente

Standardizirane komponente

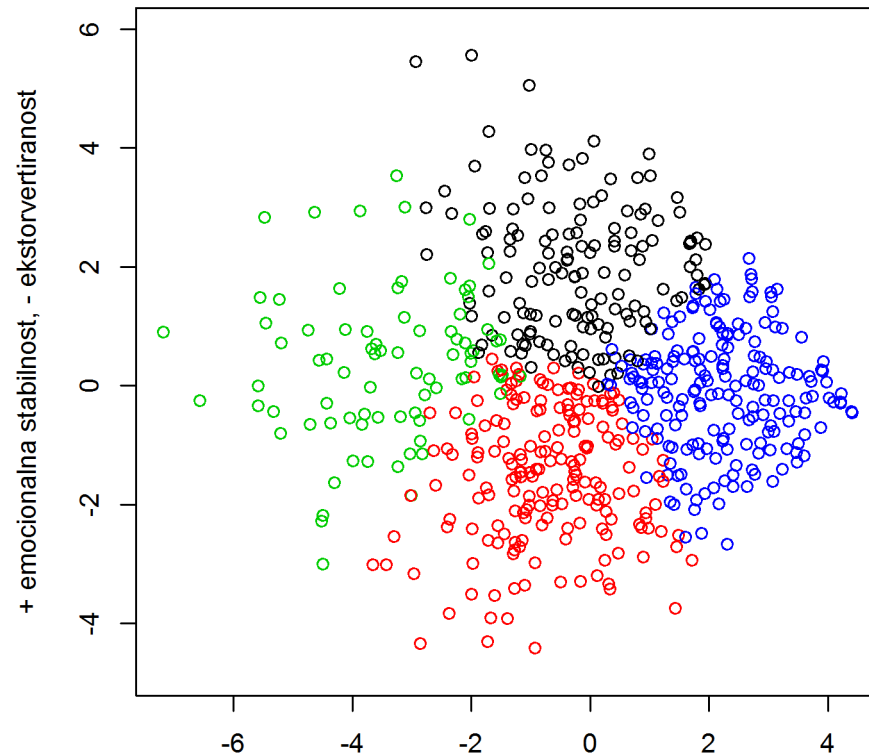
○ + emoc, - ekstr ● - emoc, + ekstr ● nizko ● visoko



Splošna komponenta (emocionalna stabilnost in ekstrverziranost)

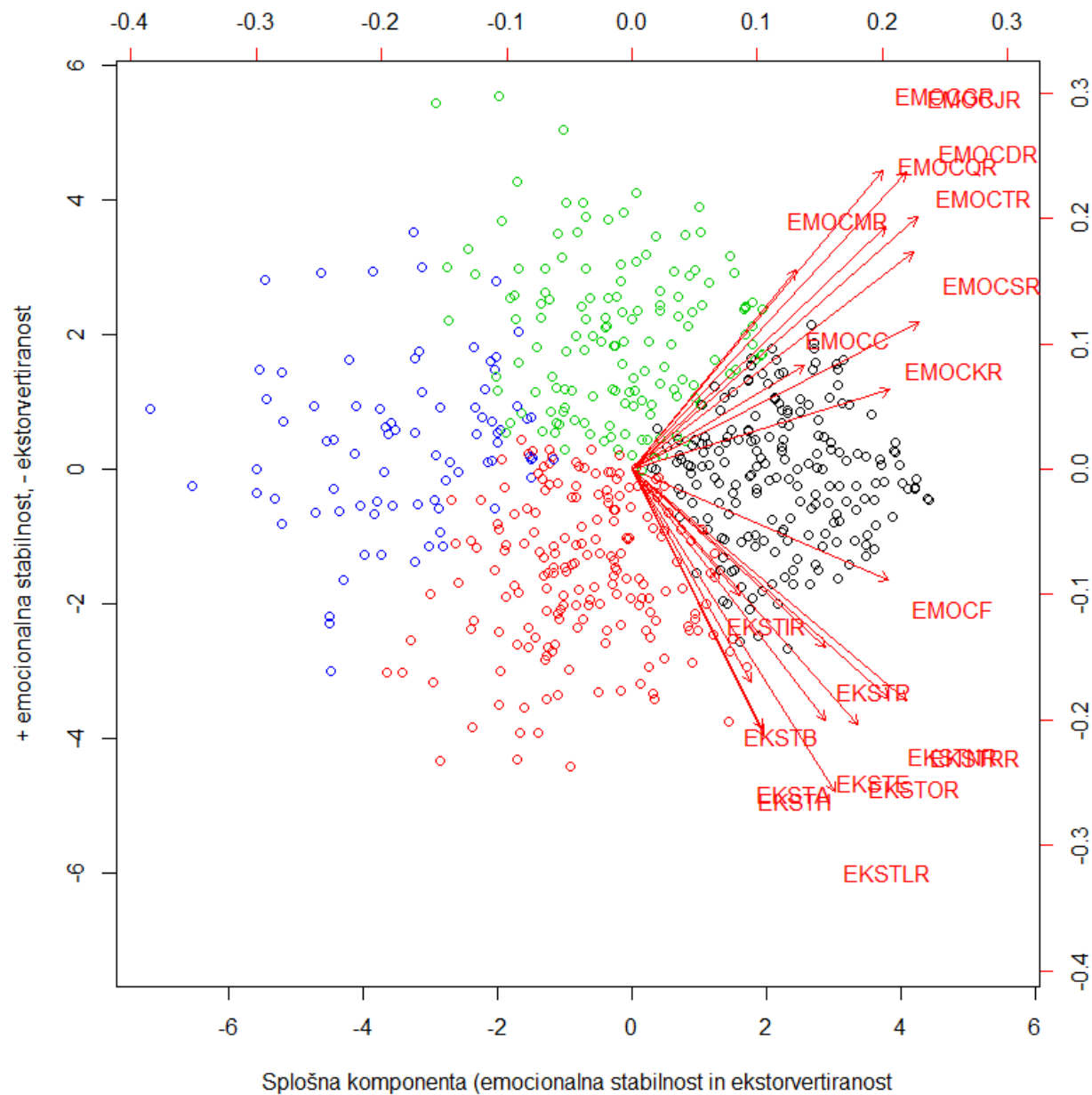
Reskalirane komponente

○ + emoc, - ekstr ● - emoc, + ekstr ● nizko ● visoko



Splošna komponenta (emocionalna stabilnost in ekstrverziranost)

Biplot

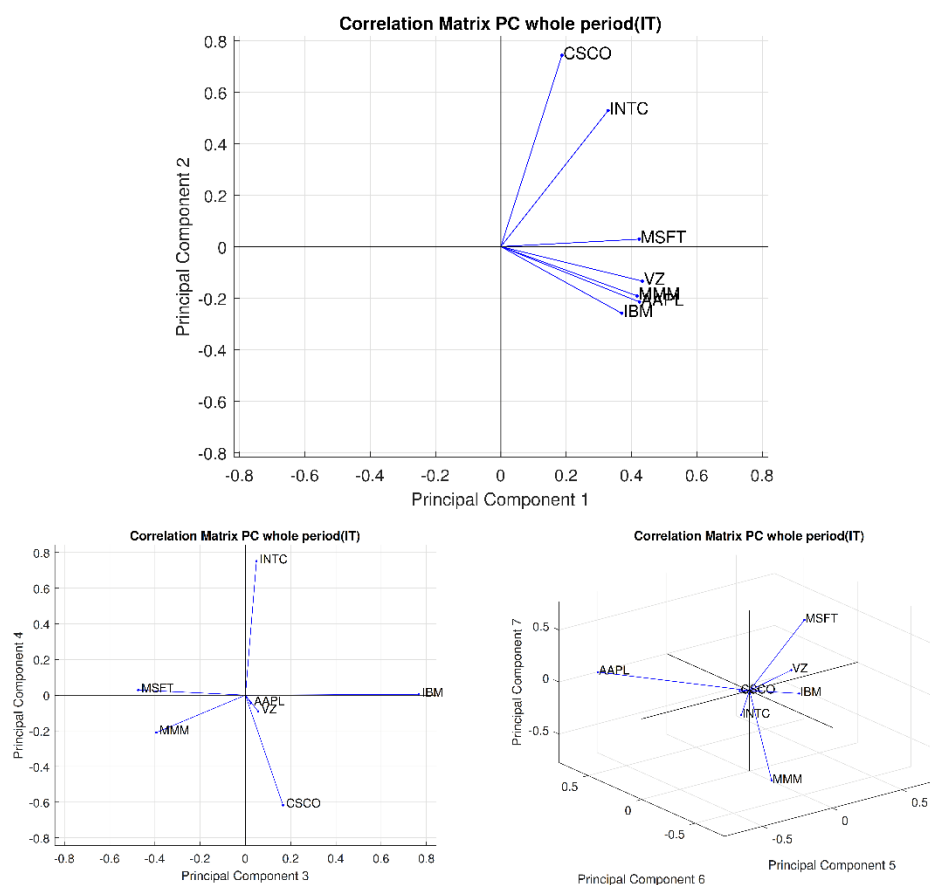


Primer iz literature

- Pasini, G. (2017). PRINCIPAL COMPONENT ANALYSIS FOR STOCK PORTFOLIO MANAGEMENT. *International Journal of Pure and Applied Mathematics*, 115(1).
<https://doi.org/10.12732/ijpam.v115i1.12>
- Vhodni podatki so korelacije med cenami delnic (spremenljivke so cene posameznih delnic, enot so dnevi)

Uporaba MGK za prikaz odnosov med cenami delnic

Figure 1: PCA application to Technology and Telecommunications Stocks



Pojasnjena variabilnost po obdobjih

Table 1: Stocks variation in different three periods

Eigenvalues and Cumulative Variance						
	Whole Period		Before Crisis		After Crisis	
PC	Eigval	Cum Var(%)	Eigval	Cum Var (%)	Eigval	Cum Var (%)
1	4.9843	71.2044	3.6394	51.9918	5.6683	80.9764
2	1.3901	91.0624	2.0329	81.0328	0.9693	94.8228
3	0.3730	96.3911	0.6278	90.0015	0.1505	96.9726
4	0.1263	98.1954	0.2866	94.0964	0.105	98.4718
5	0.0637	99.1050	0.2286	97.3620	0.0653	99.4042
6	0.0334	99.5825	0.1271	99.1782	0.0249	99.7599
7	0.0292	100	0.0575	100	0.0168	100