Stat 460 Homework 4

Wulf Novak

April 9, 2019

Factor Analysis for Flea Data

Correlation Matrix for Flea Data

	x1	x2	x3	x4
x1	1.000	0.181	-0.167	-0.074
x2	0.181	1.000	0.727	0.592
x3	-0.167	0.727	1.000	0.587
x4	-0.074	0.592	0.587	1.000

Analysis WITHOUT Varimax rotation for Flea Data

```
## [1] 0.568 0.270 0.112 0.049
```

```
## [1] 0.568 0.839 0.951 1.000
```

I have opted to choose 3 factors because this would explain over 95% of the variation, and because the first 3 factors explain at least 10% of the variation.

	Factor 1	Factor 2	Factor 3	Communalities	SpecificVar	EigValues	VarAccounted
x1	-0.039	0.989	0.047	0.983	0.017	0.568	0.568
x2	0.889	0.269	-0.203	0.903	0.097	0.270	0.839
x3	0.893	-0.157	-0.311	0.920	0.080	0.112	0.951
x4	0.827	-0.073	0.556	0.999	0.001	0.049	1.000

The first factor is primarily x2, x3, x4. The second factor is primarily x1, and the last factor is mostly composed of x4, but partly x3.

Analysis WITH Varimax rotation for Flea Data

	Factor 1	Factor 2	Factor 3	Communalities	SpecificVar	EigValues	VarAccounted
x1	0.002	0.991	-0.035	0.983	0.017	0.568	0.568
x2	0.863	0.234	0.323	0.903	0.097	0.270	0.839
x3	0.902	-0.199	0.259	0.920	0.080	0.112	0.951
x4	0.363	-0.046	0.930	0.999	0.001	0.049	1.000

Unlike without varimax rotation, the first factor primarily usues x2 and x3. The 2nd factor is primarily x1, and the 3rd factor is primarily x1. The largest difference between with varimax and without varimax is that with varimax rotation x4 is heavily weighted in factor 3, whereas without varimax rotation x4 is weighted in factor 1.

All of these commonalities are .9 or above, which suggests that these factors adequately account for the variables.

Factor Analysis for Engineer Data

Correlation Matrix for Engineer Data

	X1	X2	X3	X4	X5	X6
X1	1.000	0.189	0.101	0.123	-0.054	0.387
X2	0.189	1.000	-0.121	-0.266	-0.245	-0.067
X3	0.101	-0.121	1.000	0.289	-0.158	0.075
X4	0.123	-0.266	0.289	1.000	-0.189	0.327
X5	-0.054	-0.245	-0.158	-0.189	1.000	-0.150
X6	0.387	-0.067	0.075	0.327	-0.150	1.000

Analysis WITHOUT Varimax rotation for Engineer Data

```
## [1] 0.296 0.226 0.179 0.136 0.088 0.075
```

```
## [1] 0.296 0.522 0.700 0.836 0.925 1.000
```

I chose to use 4 factors because the first 4 factors account for at least 10% of the variation.

	Factor 1	Factor 2	Factor 3	Factor 4	Communalities	SpecificVar	EigValues	VarAccounted
X1	0.536	0.461	0.478	0.355	0.855	0.145	0.296	0.296
X2	-0.129	0.870	-0.182	0.119	0.820	0.180	0.226	0.522
X3	0.514	-0.254	-0.448	0.648	0.949	0.051	0.179	0.700
X4	0.724	-0.366	-0.110	-0.222	0.719	0.281	0.136	0.836
X5	-0.416	-0.414	0.649	0.360	0.896	0.104	0.088	0.925
X6	0.715	0.124	0.420	-0.276	0.778	0.222	0.075	1.000

Factor one is primarily x6 and x4, but also has some weight from x1 and x3. Factor 2 is primarily x2 with some weight from x1 and x5. Factor 3 is a mix of x1,x3,x5 and x6. Lastly, factor 4 is primarily x3, with small weights given to almost every other variable.

Analysis WITH Varimax rotation for Engineer Data

	Factor 1	Factor 2	Factor 3	Factor 4	Communalities	SpecificVar	EigValues	VarAccounted
X1	0.878	-0.214	0.075	0.181	0.855	0.145	0.296	0.296
X2	0.173	-0.807	-0.369	-0.054	0.820	0.180	0.226	0.522
X3	0.044	0.146	-0.090	0.958	0.949	0.051	0.179	0.700
X4	0.251	0.704	-0.323	0.238	0.719	0.281	0.136	0.836
X5	-0.024	0.044	0.941	-0.090	0.896	0.104	0.088	0.925
X6	0.758	0.379	-0.199	-0.144	0.778	0.222	0.075	1.000

With Varimax rotation that factor weightings are more dramatic. It is a lot more clear which variables are represented in each factor. Factor 1 is primarily x1 and x6. Factor 2 is composed primarily of x2 and x4. Factor 3 is mostly weighted by x5, and Factor 4 is primarily x3.

The communalities are reasonable, with x1, x2, x3, and x5 being above .82. However, x4 and x6 are below .8, which may not be preferable depending on how important it is to you to explain those specific variables.