

Nama : M.Mahbubillah
NIM : 222011569
No. Absen : 16
Kelas : 3SI1
Mata Kuliah : Analisis Peubah Ganda

Tugas Pertemuan 8

Cari tahu apa itu Bartlett's test of sphericity?

Bartlett's Test of Sphericity adalah tes yang digunakan untuk menguji in-terdependensi antara variabel-variabel yang menjadi indikator suatu faktor. Analisis ini bermaksud untuk menyatakan bahwa variabel-variabel yang dimaksud tidak berkorelasi satu dengan lainnya dalam populasi.

Asumsi Analisis Faktor yang pertama adalah: Bartlett Test of Sphericity. Rumus yang digunakan untuk **Bartlett Test of Sphericity** adalah sebagai berikut:

Bartlett Test

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

Dimana:

R = Nilai determinan; n = Jumlah data; p = jumlah variabel

Kerjakan latihan pada Buku Jhonson & Wichern (2002) exercise 8.11 atau 8.12

8_12Johnson

M.Mahbubillah

2022-10-30

```
library(readxl)
airPollution <- read_excel("airpollution.xlsx")
head(airPollution)

## # A tibble: 6 × 7
##   Wind Solar_radiation    CO    NO  NO_2  O_3    HC
##   <dbl>         <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1      8             98      7      2     12      8      2
## 2      7            107      4      3      9      5      3
## 3      7            103      4      3      5      6      3
## 4     10             88      5      2      8     15      4
```

## 5	6	91	4	2	8	10	3
## 6	8	90	5	2	12	12	4

Calculating the covariance matrix (S) and correlation matrix (R) to turn the principal component analysis and checking the eigenvalues and eigenvectors

```
S <- cov(airPollution)
print(S)
```

##	Wind	Solar_radiation	CO	NO
NO_2				
## Wind	2.5000000	-2.7804878	-0.3780488	-0.4634146
3659				
## Solar_radiation	-2.7804878	300.5156794	3.9094077	-1.3867596
0662				
## CO	-0.3780488	3.9094077	1.5220674	0.6736353
7503				
## NO	-0.4634146	-1.3867596	0.6736353	1.1823461
2695				
## NO_2	-0.5853659	6.7630662	2.3147503	1.0882695
5308				
## O_3	-2.2317073	30.7909408	2.8217189	-0.8106852
5970				
## HC	0.1707317	0.6236934	0.1416957	0.1765389
1347				
##	O_3	HC		
## Wind	-2.2317073	0.1707317		
## Solar_radiation	30.7909408	0.6236934		
## CO	2.8217189	0.1416957		
## NO	-0.8106852	0.1765389		
## NO_2	3.1265970	1.0441347		
## O_3	30.9785134	0.5946574		
## HC	0.5946574	0.4785134		

```
eigen(S)
```

```
## eigen() decomposition
## $values
## [1] 304.2578640 28.2761046 11.4644830 2.5243296 1.2795247 0.5287288
## [7] 0.2096157
##
## $vectors
##           [,1]      [,2]      [,3]      [,4]      [,5]
## [1,] 0.010039244 0.07622439 0.03087761 0.9203045748 0.3423859285
## [2,] -0.993199405 0.11615518 0.00659069 -0.0002118679 0.0022391022
## [3,] -0.014062314 -0.09956775 -0.18282641 -0.1382922410 0.6500776063
## [4,] 0.004710175 0.01320423 -0.13021553 -0.3277842624 0.6431560485
## [5,] -0.024255644 -0.15038113 -0.95526318 0.1023719020 -0.2065840405
## [6,] -0.112429558 -0.97335904 0.16981025 0.0632480276 -0.0002935726
## [7,] -0.002340785 -0.02382046 -0.08519558 0.1095073458 0.0619613872
##           [,6]      [,7]
## [1,] 0.011779079 -0.169729925
## [2,] 0.003353218 -0.001781987
## [3,] -0.563893916 0.443577538
```

```
## [4,] 0.497513370 -0.462855916
## [5,] -0.009009299 -0.105029951
## [6,] 0.051067254 -0.066992404
## [7,] 0.657012233 0.738019426
```

```
R <- cor(airPollution)
print(R)
```

```
##           Wind Solar_radiation      CO      NO
NO_2
## Wind           1.0000000    -0.10144191 -0.1938032 -0.26954261 -0.10
98249
## Solar_radiation -0.1014419    1.00000000  0.1827934 -0.07356907  0.11
57320
## CO              -0.1938032    0.18279338  1.0000000  0.50215246  0.55
65838
## NO              -0.2695426    -0.07356907  0.5021525  1.00000000  0.29
68981
## NO_2            -0.1098249    0.11573199  0.5565838  0.29689814  1.00
00000
## O_3             -0.2535928    0.31912373  0.4109288 -0.13395214  0.16
66422
## HC              0.1560979    0.05201044  0.1660323  0.23470432  0.44
77678
##           O_3      HC
## Wind          -0.2535928 0.15609793
## Solar_radiation 0.3191237 0.05201044
## CO              0.4109288 0.16603235
## NO              -0.1339521 0.23470432
## NO_2            0.1666422 0.44776780
## O_3             1.0000000 0.15445056
## HC              0.1544506 1.00000000
```

```
eigen(R)
```

```
## eigen() decomposition
```

```
## $values
```

```
## [1] 2.3367826 1.3860007 1.2040659 0.7270865 0.6534765 0.5366888 0.15589
89
```

```
##
```

```
## $vectors
```

```
##           [,1]           [,2]           [,3]           [,4]           [,5]
[,6]
## [1,] 0.2368211 0.278445138 0.6434744 0.172719491 0.56053441 -0.223
579220
## [2,] -0.2055665 -0.526613869 0.2244690 0.778136601 -0.15613432 -0.005
700851
## [3,] -0.5510839 -0.006819502 -0.1136089 0.005301798 0.57342221 -0.109
538907
## [4,] -0.3776151 0.434674253 -0.4070978 0.290503052 -0.05669070 -0.450
234781
## [5,] -0.4980161 0.199767367 0.1965567 -0.042428178 0.05021430 0.744
968707
## [6,] -0.3245506 -0.566973655 0.1598465 -0.507915905 0.08024349 -0.330
583071
```

```
## [7,] -0.3194032  0.307882771  0.5410484 -0.143082348 -0.56607057 -0.266
469812
##           [,7]
## [1,] -0.24146701
## [2,] -0.01126548
## [3,]  0.58524622
## [4,] -0.46088973
## [5,] -0.33784371
## [6,] -0.41707805
## [7,]  0.31391372
```

PCA is performed using each covariance matrix and correlation matrix

```
Spca <- princomp(x=airPollution)
print(summary(Spca),loadings = T)

## Importance of components:
##                               Comp.1      Comp.2      Comp.3      Comp.4
Comp.5
## Standard deviation      17.234083 5.25384279 3.34537279 1.569785488 1.11
7613433
## Proportion of Variance  0.872948 0.08112714 0.03289281 0.007242569 0.00
3671092
## Cumulative Proportion  0.872948 0.95407514 0.98696795 0.994210520 0.99
7881611
##                               Comp.6      Comp.7
## Standard deviation      0.718428856 0.4523547944
## Proportion of Variance  0.001516979 0.0006014096
## Cumulative Proportion  0.999398590 1.0000000000
##
## Loadings:
##               Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7
## Wind                0.920  0.342                0.170
## Solar_radiation -0.993  0.116
## CO                 -0.183 -0.138  0.650 -0.564 -0.444
## NO                 -0.130 -0.328  0.643  0.498  0.463
## NO_2                -0.150 -0.955  0.102 -0.207          0.105
## O_3                 -0.112 -0.973  0.170
## HC                  0.110          0.657 -0.738

Rpca <- princomp(x=airPollution,cor =T)
print(summary(Rpca),loadings = T)

## Importance of components:
##                               Comp.1      Comp.2      Comp.3      Comp.4      Comp
.5
## Standard deviation      1.5286539 1.1772853 1.0972994 0.8526937 0.808378
96
## Proportion of Variance  0.3338261 0.1980001 0.1720094 0.1038695 0.093353
79
## Cumulative Proportion  0.3338261 0.5318262 0.7038356 0.8077051 0.901058
89
##                               Comp.6      Comp.7
## Standard deviation      0.73259047 0.39484041
## Proportion of Variance  0.07666983 0.02227128
```

```
## Cumulative Proportion  0.97772872 1.00000000
##
## Loadings:
##           Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7
## Wind           0.237  0.278  0.643  0.173  0.561  0.224  0.241
## Solar_radiation -0.206 -0.527  0.224  0.778 -0.156
## CO              -0.551      -0.114      0.573  0.110 -0.585
## NO              -0.378  0.435 -0.407  0.291      0.450  0.461
## NO_2            -0.498  0.200  0.197      -0.745  0.338
## O_3             -0.325 -0.567  0.160 -0.508      0.331  0.417
## HC              -0.319  0.308  0.541 -0.143 -0.566  0.266 -0.314
```

PC 1 = 0.2368211X1 -0.2055665X2 -0.5510839X3 -0.3776151X4 -0.4980161X5 -0.3245506X6 -0.3194032X7

PC 2 = 0.278445138X1 -0.526613869X2 -0.006819502X3 + 0.434674253X4 + 0.199767367X5 -0.566973655X6 + 0.307882771X7

PC 3 = 0.6434744X1 + 0.2244690X2 -0.1136089X3 -0.4070978X4 + 0.1965567X5 +0.1598465X6 + 0.5410484X7

PC 4 = 0.172719491X1 + 0.778136601X2 +0.005301798X3 +0.290503052X4 -0.042428178X5 -0.507915905X6 -0.143082348X7

Interpretasi :

- Hasil dari penggunaan kovarians (S), menunjukkan bahwa komponen utama pertama dapat menjelaskan 87% dari data. Variabel utama dari komponen utama pertama adalah Solar_radiation

- Ketika matriks korelasi (R) digunakan, 80% data dapat dijelaskan dengan mengumpulkan komponen utama pertama, kedua, ketiga, dan keempat. Komponen utamanya terdiri dari berbagai elemen.

- Jika skala (unit) setiap variabel berbeda dalam kumpulan data, hasil yang baik dapat diperoleh dengan menormalkan kovarians atau menggunakan matriks korelasi. Oleh karena itu, dinilai lebih tepat menggunakan korelasi ketika membandingkan variabel dengan skala yang berbeda.

- Ketika analisis komponen utama menggunakan matriks korelasi dilakukan, dinilai lebih baik menggunakan setidaknya 4 komponen utama. Dengan kata lain, jika komponen utama pertama, kedua, ketiga, dan keempat digunakan, maka 80% dari total data dapat dijelaskan, sehingga dapat dikatakan bahwa daya penjas cukup.