

Practical No 7

Aim: Demonstration of Principal Component Analysis.

Theory

Steps:

Step1: click on packages and set cran mirror(click on other and select USA IN)

Step2: click on packages and select install packages and install package FactoMineR.

install.packages("FactoMineR")

library(FactoMineR)

Step3: Create Excel Sheet.

	A	B	C	
	Math	English	Art	
1	90	40	40	
2	80	55	50	
3	70	35	60	
4	60	68	70	
5	50	78	80	
6				
7				

Code:

```
x=read.csv("C:/Users/Akki/OneDrive/Desktop/JEENAL/students.csv")
```

```
x
```

```
> x=read.csv("C:/Users/Akki/OneDrive/Desktop/JEENAL/students.csv")
```

```
> x
```

```
  Math English Art
```

```
1   90      40  40
```

```
2   80      55  50
```

```
3   70      35  60
```

```
4   60      68  70
```

```
5   50      78  80
```

```
cov_mat=cov(x)
```

```
cov_mat
```

```
> cov_mat=cov(x)
```

```
> cov_mat
```

```
      Math English  Art
```

```
Math    250.0  -222.5 -250.0
```

```
English -222.5   330.7  222.5
```

```
Art     -250.0   222.5  250.0
```

```
ex=eigen(cov_mat)
```

```
ex
```

```

> ex=eigen(cov_mat)
> ex
eigen() decomposition
$values
[1] 7.411998e+02 8.950015e+01 2.991369e-14

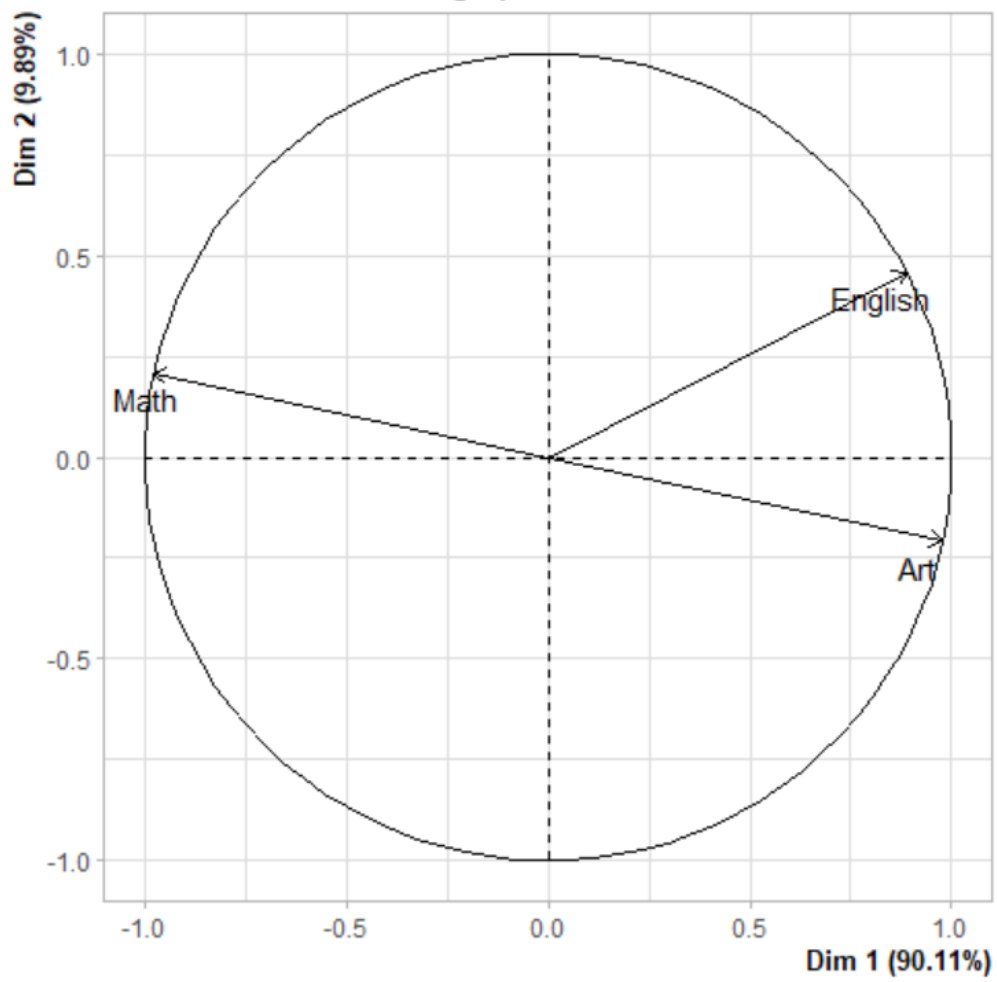
$vectors
      [,1]      [,2]      [,3]
[1,]  0.5612001 -0.4301795 7.071068e-01
[2,] -0.6083657 -0.7936568 2.220446e-16
[3,] -0.5612001  0.4301795 7.071068e-01

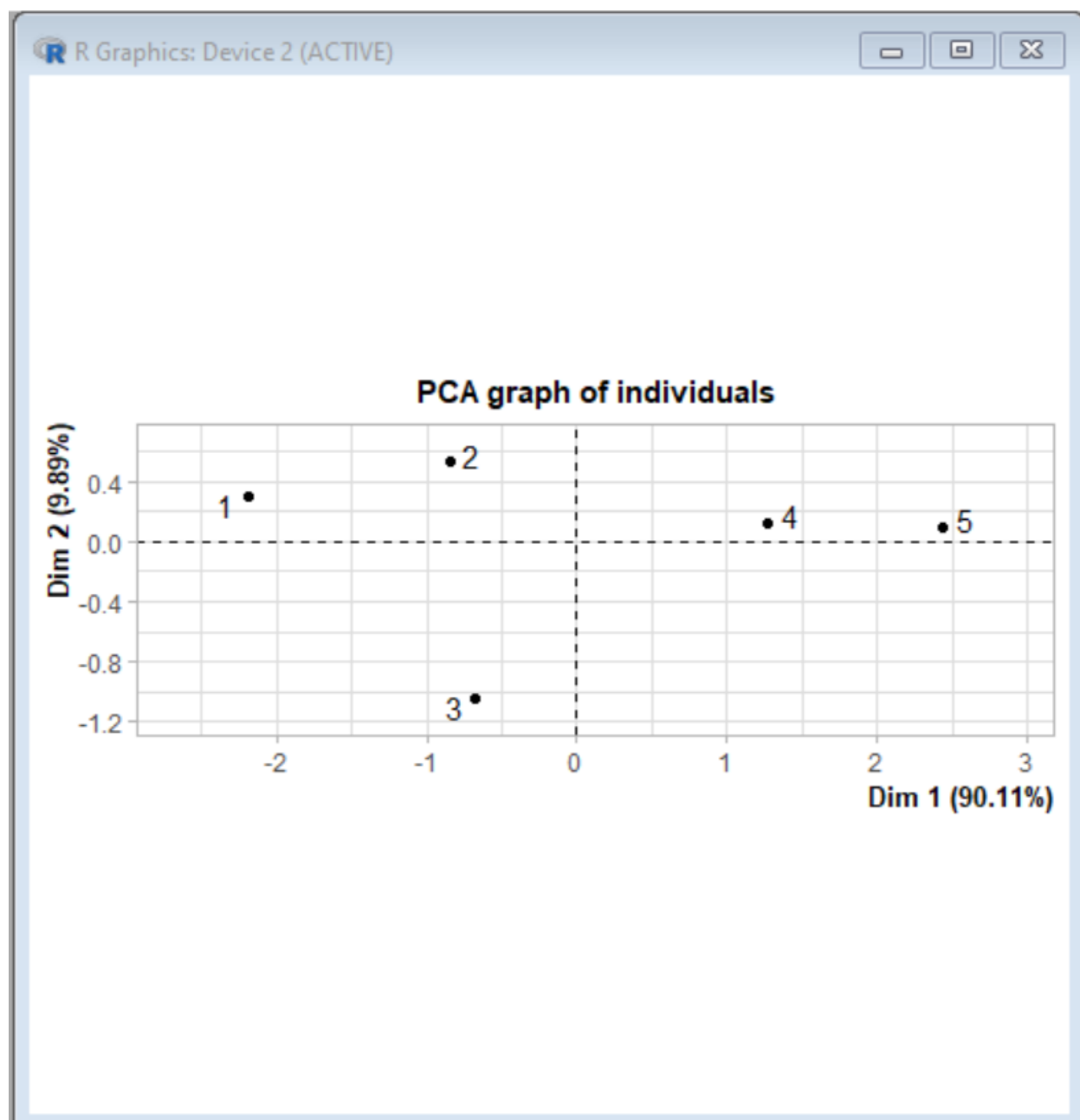
> datapca=PCA(x,ncp=3,graph=TRUE)

```

```
datapca=PCA(x,ncp=3,graph=TRUE)
```

PCA graph of variables





```
datapca$eig
```

```
> datapca$eig
```

```
      eigenvalue percentage of variance cumulative percentage of variance
comp 1  2.7031671          90.105571          90.10557
comp 2  0.2968329           9.894429         100.00000
comp 3  0.0000000           0.000000         100.00000
```

```
datapca$var
```

```
> datapca$var
$coord
      Dim.1      Dim.2 Dim.3
Math    -0.9780749   0.2082535    0
English  0.8887667   0.4583599    0
Art      0.9780749  -0.2082535    0

$cor
      Dim.1      Dim.2 Dim.3
Math    -0.9780749   0.2082535    0
English  0.8887667   0.4583599    0
Art      0.9780749  -0.2082535    0

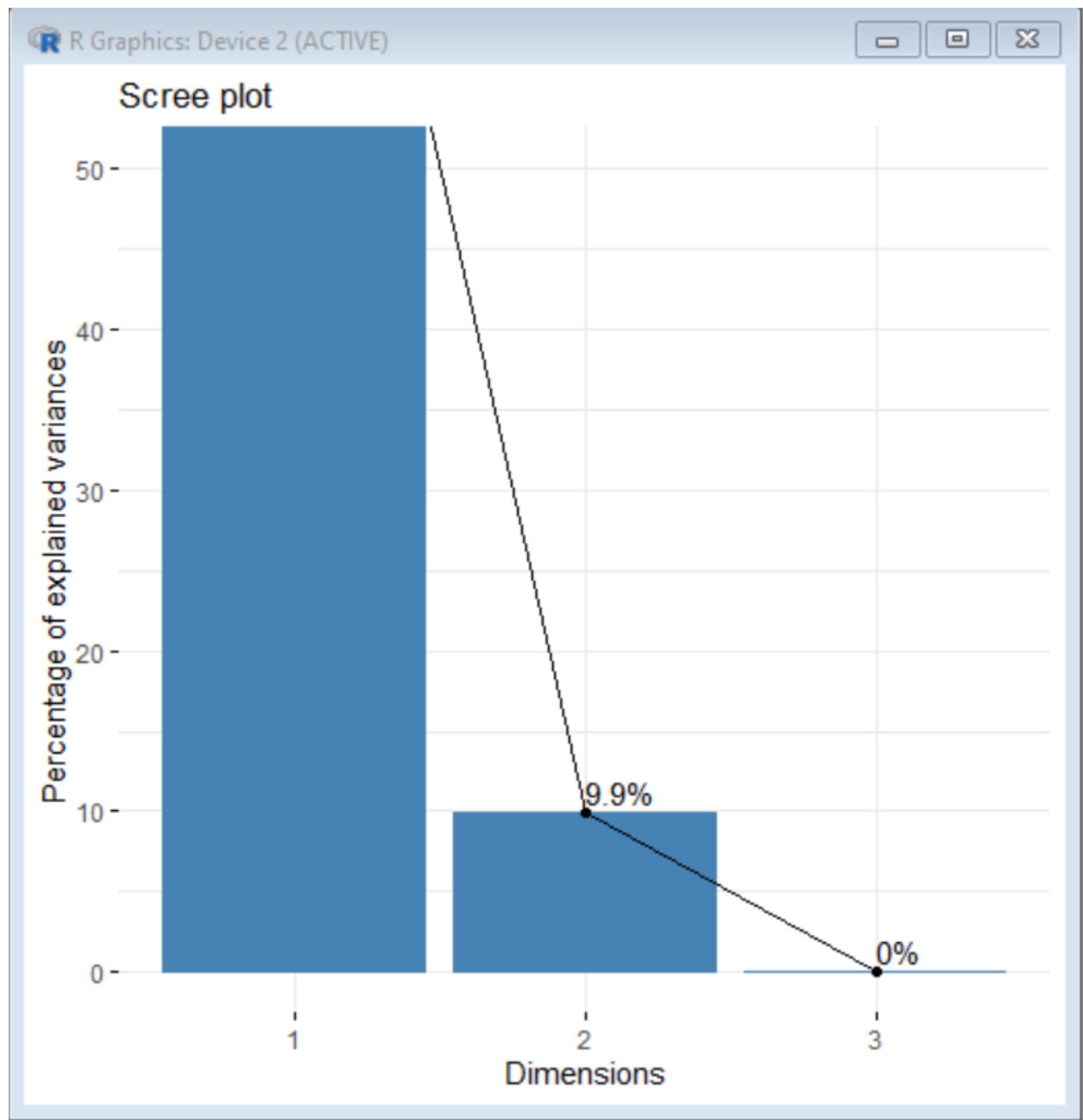
$cos2
      Dim.1      Dim.2 Dim.3
Math    0.9566305  0.04336952    0
English 0.7899062  0.21009384    0
Art      0.9566305  0.04336952    0

$contrib
      Dim.1      Dim.2 Dim.3
Math    35.38925  14.61075   NaN
English 29.22151  70.77849   NaN
Art      35.38925  14.61075   NaN
```

datapca\$var\$coord

```
> datapca$var$coord
      Dim.1      Dim.2 Dim.3
Math    -0.9780749   0.2082535    0
English  0.8887667   0.4583599    0
Art      0.9780749  -0.2082535    0
> |
```

fviz_screplot(datapca,addlabels=TRUE,ylim=c(0,50))



```
install.packages('factoextra',repos="http://cran.us.r-project.org")
```

```
> install.packages('factoextra', repos="http://cran.us.r-project.org")
Installing package into 'C:/Users/Akki/AppData/Local/R/win-library/4.2'
(as 'lib' is unspecified)
also installing the dependencies 'corrplot', 'viridis', 'ggsci', 'cowplot', 'gg$

trying URL 'http://cran.us.r-project.org/bin/windows/contrib/4.2/corrplot_0.92.$
Content type 'application/zip' length 3844780 bytes (3.7 MB)
downloaded 3.7 MB

trying URL 'http://cran.us.r-project.org/bin/windows/contrib/4.2/viridis_0.6.2.$
Content type 'application/zip' length 2999945 bytes (2.9 MB)
downloaded 2.9 MB

trying URL 'http://cran.us.r-project.org/bin/windows/contrib/4.2/ggsci_2.9.zip'
Content type 'application/zip' length 2978453 bytes (2.8 MB)
downloaded 2.8 MB

trying URL 'http://cran.us.r-project.org/bin/windows/contrib/4.2/cowplot_1.1.1.$
```

```
library("factoextra")
fviz_screplot(datapca, addlabels=TRUE, ylim=c(0,50))
head(iris)
x=iris[,-5]
```

```
> library("factoextra")
Loading required package: ggplot2
Learn more about the underlying theory at https://ggplot2-book.org/
Welcome! Want to learn more? See two factoextra-related books at https://go
> fviz_screplot(datapca, addlabels=TRUE, ylim=c(0,50))
> head(iris)
  Sepal.Length Sepal.Width Petal.Length Petal.Width Species
1          5.1         3.5         1.4         0.2  setosa
2          4.9         3.0         1.4         0.2  setosa
3          4.7         3.2         1.3         0.2  setosa
4          4.6         3.1         1.5         0.2  setosa
5          5.0         3.6         1.4         0.2  setosa
6          5.4         3.9         1.7         0.4  setosa
> xiri[,-5]
Error: object 'xiri' not found
> x=iris[,-5]
```

```
x=iris[,-5]
x
```

```
> x=iris[,-5]
> x
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
1	5.1	3.5	1.4	0.2
2	4.9	3.0	1.4	0.2
3	4.7	3.2	1.3	0.2
4	4.6	3.1	1.5	0.2
5	5.0	3.6	1.4	0.2
6	5.4	3.9	1.7	0.4
7	4.6	3.4	1.4	0.3
8	5.0	3.4	1.5	0.2
9	4.4	2.9	1.4	0.2
10	4.9	3.1	1.5	0.1
11	5.4	3.7	1.5	0.2
12	4.8	3.4	1.6	0.2
13	4.8	3.0	1.4	0.1
14	4.3	3.0	1.1	0.1
15	5.8	4.0	1.2	0.2

```
cov_iris=cov(x)
cov_iris
```

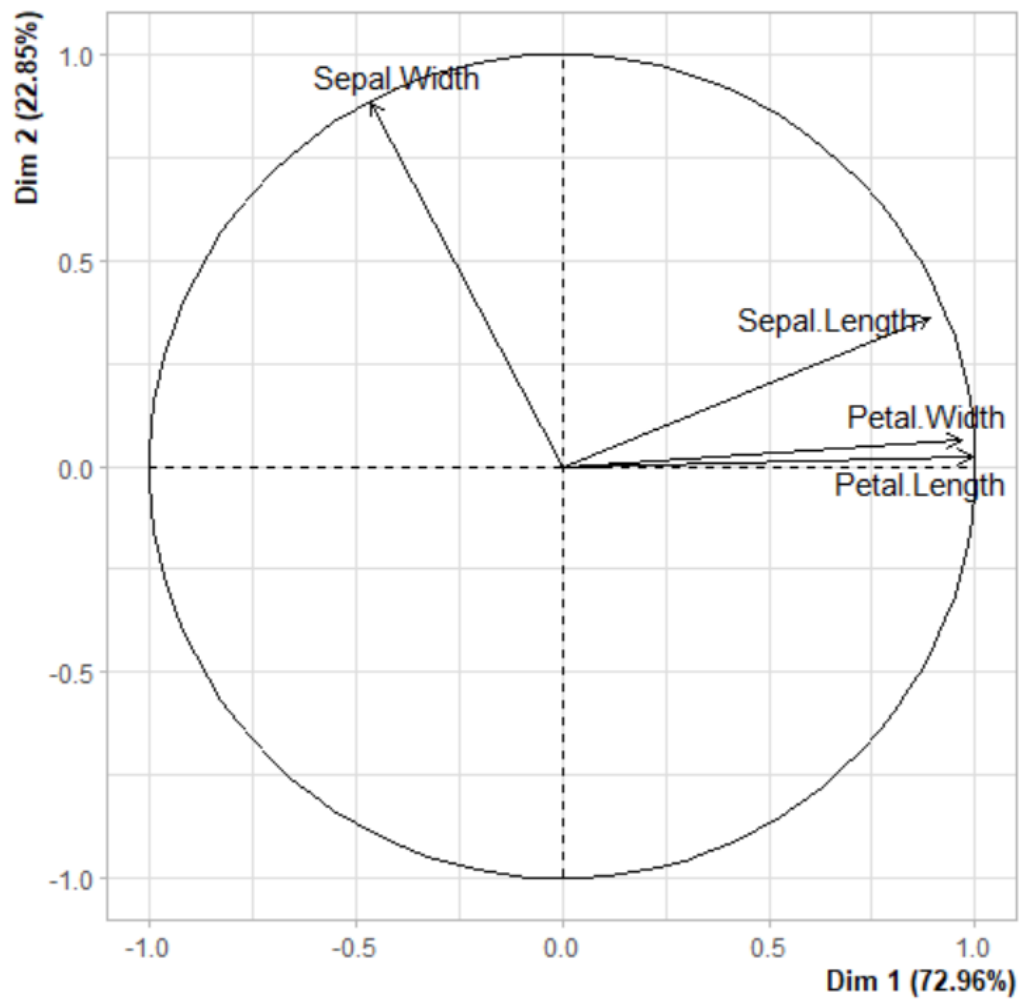
```
> cov_iris=cov(x)
> cov_iris
```

	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
Sepal.Length	0.6856935	-0.0424340	1.2743154	0.5162707
Sepal.Width	-0.0424340	0.1899794	-0.3296564	-0.1216394
Petal.Length	1.2743154	-0.3296564	3.1162779	1.2956094
Petal.Width	0.5162707	-0.1216394	1.2956094	0.5810063

```
> irispca=PCA(x,ncp=3,graph=TRUE)
```

```
> irispca=PCA(x,ncp=3,graph=TRUE)
```


PCA graph of variables





```
> irispca
**Results for the Principal Component Analysis (PCA)**
The analysis was performed on 150 individuals, described by 4 variables
*The results are available in the following objects:

      name          description
1  "$eig"          "eigenvalues"
2  "$var"          "results for the variables"
3  "$var$coord"    "coord. for the variables"
4  "$var$cor"      "correlations variables - dimensions"
5  "$var$cos2"     "cos2 for the variables"
6  "$var$contrib"  "contributions of the variables"
7  "$ind"          "results for the individuals"
8  "$ind$coord"    "coord. for the individuals"
9  "$ind$cos2"     "cos2 for the individuals"
10 "$ind$contrib"  "contributions of the individuals"
11 "$call"         "summary statistics"
12 "$call$centre"  "mean of the variables"
13 "$call$ecart.type" "standard error of the variables"
14 "$call$row.w"   "weights for the individuals"
15 "$call$col.w"   "weights for the variables"
> |
```

summary(irispca)

```
> summary(irispca)

Call:
PCA(X = x, ncp = 3, graph = TRUE)

Eigenvalues

          Dim.1   Dim.2   Dim.3   Dim.4
Variance      2.918   0.914   0.147   0.021
% of var.     72.962  22.851   3.669   0.518
Cumulative % of var. 72.962  95.813  99.482 100.000

Individuals (the 10 first)

      Dist   Dim.1   ctr   cos2   Dim.2   ctr   cos2   Dim.3
1      | 2.319 | -2.265 1.172 0.954 | 0.480 0.168 0.043 | -0.128
2      | 2.202 | -2.081 0.989 0.893 | -0.674 0.331 0.094 | -0.235
3      | 2.389 | -2.364 1.277 0.979 | -0.342 0.085 0.020 | 0.044
4      | 2.378 | -2.299 1.208 0.935 | -0.597 0.260 0.063 | 0.091
5      | 2.476 | -2.390 1.305 0.932 | 0.647 0.305 0.068 | 0.016
6      | 2.555 | -2.076 0.984 0.660 | 1.489 1.617 0.340 | 0.027
7      | 2.468 | -2.444 1.364 0.981 | 0.048 0.002 0.000 | 0.335
8      | 2.246 | -2.233 1.139 0.988 | 0.223 0.036 0.010 | -0.089
9      | 2.592 | -2.335 1.245 0.812 | -1.115 0.907 0.185 | 0.145
```