

Métodos Quantitativos

Prof. Dr. A. L. Korzenowski

Aula 05: Análise Fatorial Exploratória

Quando falamos de análises onde buscamos estabelecer relações de causa e efeito ou que interferem uma no comportamento da outra, consideramos ao menos uma variável independente que tenta explicar o comportamento de uma variável dependente. Em análise multivariada, tem-se um conjunto de variáveis independentes e deseja-se a construção de índices ou indicadores que representem fatores ocultos. Estes fatores chamam-se variáveis latentes e usualmente representam constructos teóricos que norteiam a escolha das variáveis escolhidas.

O objetivo da análise fatorial é o redimensionamento da base de dados de um número v de variáveis mensuráveis para um número k ($k \ll v$) de variáveis latentes. Em geral são necessárias grandes amostras para a execução destes procedimentos. A literatura recomenda pelo menos 10 casos por variável mensurada.

O procedimento possui pressupostos básicos que necessitam ser verificados:

- Variáveis correlacionadas ($r > 0,30$);
- Teste de esfericidade de Bartlett significativo (verifica se a matriz de correlações difere de uma matriz identidade);
- Adequação da amostra ($KMO > 0,60$).

Existem vários métodos de extração de fatores, sendo os mais usuais

- *Principal components*: Avalia a variância total;
- *Axis factoring*: Avalia apenas a variância comum através das comunalidades (medida da variância compartilhada das variáveis)

A retenção dos fatores pode ser por diferentes regras: Autovalor > 1 , número de fatores pré-definido, percentagem de variância explicada pré-definida e Scree-plot são os mais usuais. A fim de melhorar a interpretação dos fatores, pode-se aplicar uma rotação na solução encontrada:

- Quatimax: cria um super fator inicial
- Varimax: distribui as importâncias dos fatores e concentra as cargas das variáveis
- Equimax: combinação dos anteriores
- Oblíquos: geram rotações sem respeitar a ortogonalidade dos fatores

A cada variável é definido um peso “de participação” em cada fator. As variáveis com maiores pesos ditam o significado do fator. Em geral deseja-se cargas fatoriais $> 0,4$. A tabela apresenta os limites de significância estatística para as cargas fatoriais em função do tamanho da amostra com um poder de 80% e nível de significância de 5%.

A questão fundamental da análise é

- Atribuir um nome descritivo aos fatores retidos
- Detectar a essência das variáveis individuais
- Abstrair o objeto de pesquisa

Carga Fatorial	Tamanho Mínimo de Amostra
0,30	350
0,35	250
0,40	200
0,45	150
0,50	120
0,55	100
0,60	85
0,65	70
0,70	60
0,75	50

Componentes Principais

A função `princomp()` produz uma análise de componentes principais sem rotação. Considere a base de dados *USArrests*.

```
# Principal Components Analysis
# entering raw data and extracting PCs
# from the correlation matrix
attach(USArrests)
head(USArrests)
```

```
##           Murder Assault UrbanPop Rape
## Alabama      13.2      236      58 21.2
## Alaska       10.0      263      48 44.5
## Arizona        8.1      294      80 31.0
## Arkansas       8.8      190      50 19.5
## California     9.0      276      91 40.6
## Colorado       7.9      204      78 38.7
```

```
mydata <- USArrests
summary(mydata)
```

```
##           Murder           Assault           UrbanPop           Rape
## Min.      : 0.800   Min.      : 45.0   Min.      :32.00   Min.      : 7.30
## 1st Qu.: 4.075   1st Qu.:109.0   1st Qu.:54.50   1st Qu.:15.07
## Median : 7.250   Median :159.0   Median :66.00   Median :20.10
## Mean      : 7.788   Mean      :170.8   Mean      :65.54   Mean      :21.23
## 3rd Qu.:11.250   3rd Qu.:249.0   3rd Qu.:77.75   3rd Qu.:26.18
## Max.      :17.400   Max.      :337.0   Max.      :91.00   Max.      :46.00
```

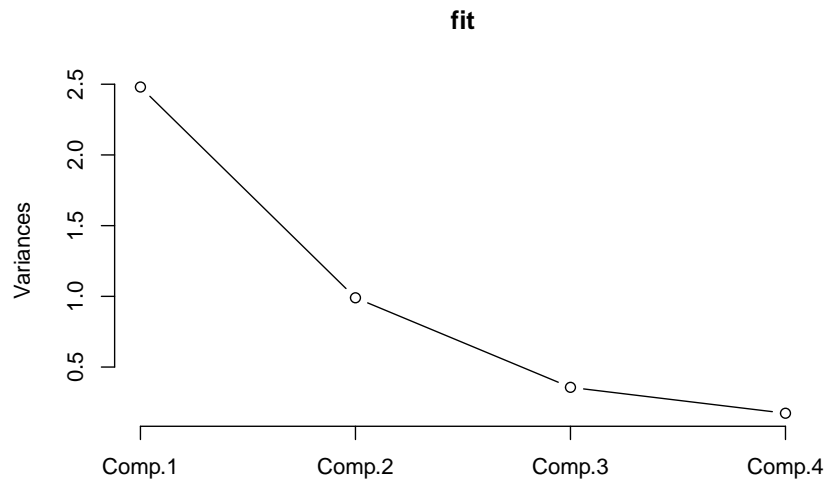
```
fit <- princomp(mydata, cor=TRUE)
summary(fit) # print variance accounted for
```

```
## Importance of components:
##              Comp.1   Comp.2   Comp.3   Comp.4
## Standard deviation  1.5748783 0.9948694 0.5971291 0.41644938
## Proportion of Variance 0.6200604 0.2474413 0.0891408 0.04335752
## Cumulative Proportion 0.6200604 0.8675017 0.9566425 1.00000000
```

```
loadings(fit) # pc loadings
```

```
##
## Loadings:
##      Comp.1 Comp.2 Comp.3 Comp.4
## Murder    0.536  0.418  0.341  0.649
## Assault    0.583  0.188  0.268 -0.743
## UrbanPop    0.278 -0.873  0.378  0.134
## Rape       0.543 -0.167 -0.818
##
##      Comp.1 Comp.2 Comp.3 Comp.4
## SS loadings    1.00  1.00  1.00  1.00
## Proportion Var  0.25  0.25  0.25  0.25
## Cumulative Var  0.25  0.50  0.75  1.00
```

```
plot(fit,type="lines") # scree plot
```

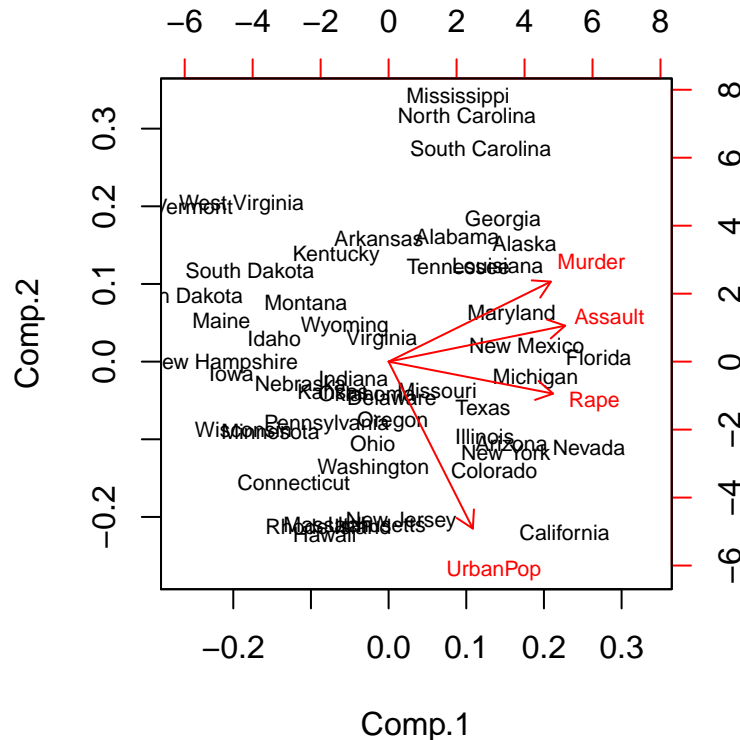


```
fit$scores # the principal components
```

```
##      Comp.1      Comp.2      Comp.3      Comp.4
## Alabama    0.98556588  1.13339238  0.44426879  0.156267145
## Alaska     1.95013775  1.07321326 -2.04000333 -0.438583440
## Arizona     1.76316354 -0.74595678 -0.05478082 -0.834652924
## Arkansas   -0.14142029  1.11979678 -0.11457369 -0.182810896
## California  2.52398013 -1.54293399 -0.59855680 -0.341996478
## Colorado    1.51456286 -0.98755509 -1.09500699  0.001464887
## Connecticut -1.35864746 -1.08892789  0.64325757 -0.118469414
## Delaware    0.04770931 -0.32535892  0.71863294 -0.881977637
## Florida     3.01304227  0.03922851  0.57682949 -0.096284752
## Georgia     1.63928304  1.27894240  0.34246008  1.076796812
## Hawaii     -0.91265715 -1.57046001 -0.05078189  0.902806864
## Idaho      -1.63979985  0.21097292 -0.25980134 -0.499104101
## Illinois    1.37891072 -0.68184119  0.67749564 -0.122021292
## Indiana    -0.50546136 -0.15156254 -0.22805484  0.424665700
## Iowa       -2.25364607 -0.10405407 -0.16456432  0.017555916
## Kansas     -0.79688112 -0.27016470 -0.02555331  0.206496428
```

## Kentucky	-0.75085907	0.95844029	0.02836942	0.670556671
## Louisiana	1.56481798	0.87105466	0.78348036	0.454728038
## Maine	-2.39682949	0.37639158	0.06568239	-0.330459817
## Maryland	1.76336939	0.42765519	0.15725013	-0.559069521
## Massachusetts	-0.48616629	-1.47449650	0.60949748	-0.179598963
## Michigan	2.10844115	-0.15539682	-0.38486858	0.102372019
## Minnesota	-1.69268181	-0.63226125	-0.15307043	0.067316885
## Mississippi	0.99649446	2.39379599	0.74080840	0.215508013
## Missouri	0.69678733	-0.26335479	-0.37744383	0.225824461
## Montana	-1.18545191	0.53687437	-0.24688932	0.123742227
## Nebraska	-1.26563654	-0.19395373	-0.17557391	0.015892888
## Nevada	2.87439454	-0.77560020	-1.16338049	0.314515476
## New Hampshire	-2.38391541	-0.01808229	-0.03685539	-0.033137338
## New Jersey	0.18156611	-1.44950571	0.76445355	0.243382700
## New Mexico	1.98002375	0.14284878	-0.18369218	-0.339533597
## New York	1.68257738	-0.82318414	0.64307509	-0.013484369
## North Carolina	1.12337861	2.22800338	0.86357179	-0.954381667
## North Dakota	-2.99222562	0.59911882	-0.30127728	-0.253987327
## Ohio	-0.22596542	-0.74223824	0.03113912	0.473915911
## Oklahoma	-0.31178286	-0.28785421	0.01530979	0.010332321
## Oregon	0.05912208	-0.54141145	-0.93983298	-0.237780688
## Pennsylvania	-0.88841582	-0.57110035	0.40062871	0.359061124
## Rhode Island	-0.86377206	-1.49197842	1.36994570	-0.613569430
## South Carolina	1.32072380	1.93340466	0.30053779	-0.131466685
## South Dakota	-1.98777484	0.82334324	-0.38929333	-0.109571764
## Tennessee	0.99974168	0.86025130	-0.18808295	0.652864291
## Texas	1.35513821	-0.41248082	0.49206886	0.643195491
## Utah	-0.55056526	-1.47150461	-0.29372804	-0.082314047
## Vermont	-2.80141174	1.40228806	-0.84126309	-0.144889914
## Virginia	-0.09633491	0.19973529	-0.01171254	0.211370813
## Washington	-0.21690338	-0.97012418	-0.62487094	-0.220847793
## West Virginia	-2.10858541	1.42484670	-0.10477467	0.131908831
## Wisconsin	-2.07971417	-0.61126862	0.13886500	0.184103743
## Wyoming	-0.62942666	0.32101297	0.24065923	-0.166651801

```
biplot(fit, cex=0.65)
```



Use **cor=FALSE** para que as componentes principais sejam extraídas a partir da matriz de covariâncias. Use a opção **covmat=** para entrar com a matriz de correlação ou covariância diretamente. Se entrar com a matriz de covariâncias, a opção **n.obs=** é obrigatória.

A função **principal()** do pacote **psych** pode ser utilizada para extrair e rotar os fatores pelo método das componentes principais.

```
# Varimax Rotated Principal Components
# retaining 5 components
library(psych)
fit <- principal(mydata, nfactors=4, rotate="varimax")
fit # print results
```

```
## Principal Components Analysis
## Call: principal(r = mydata, nfactors = 4, rotate = "varimax")
## Standardized loadings (pattern matrix) based upon correlation matrix
##          RC1  RC2  RC3  RC4  h2      u2 com
## Murder    0.91 -0.01 0.26 0.31  1  3.3e-16 1.4
## Assault   0.52  0.13 0.33 0.78  1 -2.2e-16 2.2
## UrbanPop  0.01  0.98 0.18 0.08  1  3.0e-15 1.1
## Rape      0.28  0.23 0.90 0.25  1  1.2e-15 1.5
##
##          RC1  RC2  RC3  RC4
## SS loadings      1.18 1.03 1.02 0.77
## Proportion Var    0.30 0.26 0.25 0.19
## Cumulative Var    0.30 0.55 0.81 1.00
## Proportion Explained 0.30 0.26 0.25 0.19
## Cumulative Proportion 0.30 0.55 0.81 1.00
##
## Mean item complexity = 1.6
```

```
## Test of the hypothesis that 4 components are sufficient.
##
## The root mean square of the residuals (RMSR) is 0
## with the empirical chi square 0 with prob < NA
##
## Fit based upon off diagonal values = 1
```

mydata pode ser a matriz de dados brutos ou a matriz de covariâncias. Exclusão pareada (*Pairwise deletion*) dos dados faltantes é utilizado e o tipo de rotação pode ser *none*, *varimax*, *quatimax*, *promax*, *oblimin*, *simplimax*, ou *cluster*.

Análise Fatorial Exploratória

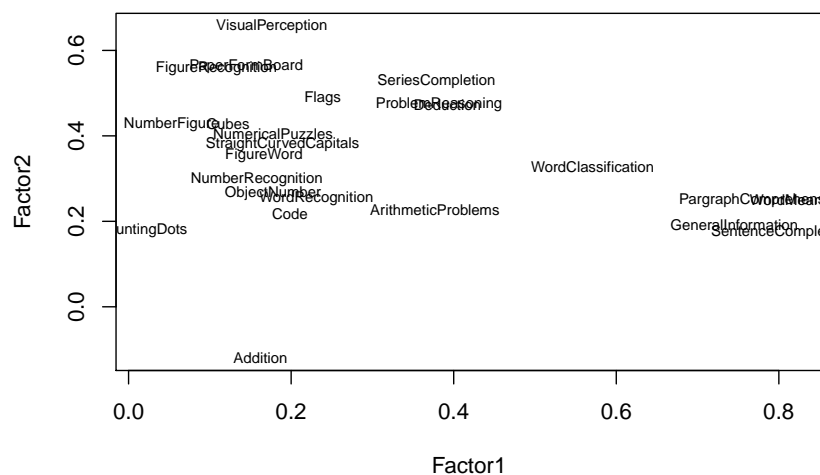
A função **factanal()** produz a análise fatorial pela função de máxima verossimilhança. Vamos considerar a matriz de correlação de 24 testes psicológicos aplicados em 145 alunos desétima e oitava séries num suburbio de Chicago por Holzinger e Swineford.

```
attach(Harman74.cor)
mydata <- Harman74.cor$cov
# Maximum Likelihood Factor Analysis
# entering raw data and extracting 3 factors,
# with varimax rotation
fit <- factanal(factors = 3, covmat = Harman74.cor, rotation = "varimax")
print(fit, digits=2, cutoff=.3, sort=TRUE)
```

```
##
## Call:
## factanal(factors = 3, covmat = Harman74.cor, rotation = "varimax")
##
## Uniquenesses:
##      VisualPerception      Cubes      PaperFormBoard
##              0.50              0.79              0.66
##           Flags      GeneralInformation      PargraphComprehension
##              0.69              0.35              0.32
##      SentenceCompletion      WordClassification      WordMeaning
##              0.30              0.50              0.26
##           Addition      Code      CountingDots
##              0.20              0.59              0.49
## StraightCurvedCapitals      WordRecognition      NumberRecognition
##              0.57              0.84              0.85
##      FigureRecognition      ObjectNumber      NumberFigure
##              0.64              0.78              0.64
##           FigureWord      Deduction      NumericalPuzzles
##              0.79              0.59              0.58
##      ProblemReasoning      SeriesCompletion      ArithmeticProblems
##              0.60              0.50              0.50
##
## Loadings:
##              Factor1 Factor2 Factor3
## GeneralInformation      0.75
## PargraphComprehension      0.78
## SentenceCompletion      0.80
```

```
## WordClassification      0.57    0.33
## WordMeaning             0.82
## VisualPerception        0.66
## PaperFormBoard          0.56
## FigureRecognition        0.56
## SeriesCompletion         0.38    0.53
## Addition                0.87
## Code                    0.57
## CountingDots            0.69
## ArithmeticProblems      0.38    0.55
## Cubes                   0.43
## Flags                   0.49
## StraightCurvedCapitals 0.38    0.50
## WordRecognition
## NumberRecognition
## ObjectNumber            0.34
## NumberFigure            0.43    0.42
## FigureWord              0.35
## Deduction               0.39    0.47
## NumericalPuzzles        0.41    0.47
## ProblemReasoning        0.38    0.47
##
##               Factor1 Factor2 Factor3
## SS loadings    3.80    3.49    3.19
## Proportion Var 0.16    0.15    0.13
## Cumulative Var 0.16    0.30    0.44
##
## Test of the hypothesis that 3 factors are sufficient.
## The chi square statistic is 295.59 on 207 degrees of freedom.
## The p-value is 5.12e-05
```

```
# plot factor 1 by factor 2
load <- fit$loadings[,1:2]
plot(load, type="n") # set up plot
text(load[,1], load[,2], labels=rownames(load), cex=0.7) # add variable names
```



A opção **rotation=** inclui as opções *varimax*, *promax*, e *none*. A opção **scores=** *regression* ou *Bartlett* determina como serão produzidos os escores. Use a opção **covmat=** para entrar com a matriz de correlação ou covariância diretamente. Se entrar com a matriz de covariâncias, a opção **n.obs=** é obrigatória.

A função `fa()` do pacote **psych** oferece uma série de funções relacionadas a análise fatorial, incluindo *principal axis factoring*.

```
# Principal Axis Factor Analysis
```

```
library(psych)
```

```
fit <- fa(mydata, nfactors=3, rotate="varimax")
```

```
fit # print results
```

```
## Factor Analysis using method = minres
## Call: fa(r = mydata, nfactors = 3, rotate = "varimax")
## Standardized loadings (pattern matrix) based upon correlation matrix
##
```

	MR1	MR3	MR2	h2	u2	com
## VisualPerception	0.17	0.65	0.21	0.49	0.51	1.4
## Cubes	0.12	0.43	0.09	0.21	0.79	1.2
## PaperFormBoard	0.16	0.55	0.01	0.33	0.67	1.2
## Flags	0.24	0.49	0.09	0.31	0.69	1.5
## GeneralInformation	0.74	0.19	0.24	0.64	0.36	1.4
## ParagraphComprehension	0.77	0.25	0.14	0.67	0.33	1.3
## SentenceCompletion	0.82	0.17	0.16	0.72	0.28	1.2
## WordClassification	0.57	0.33	0.25	0.50	0.50	2.0
## WordMeaning	0.82	0.24	0.11	0.74	0.26	1.2
## Addition	0.17	-0.13	0.83	0.74	0.26	1.1
## Code	0.18	0.18	0.63	0.46	0.54	1.3
## CountingDots	0.03	0.15	0.67	0.47	0.53	1.1
## StraightCurvedCapitals	0.19	0.35	0.51	0.42	0.58	2.1
## WordRecognition	0.22	0.24	0.26	0.17	0.83	2.9
## NumberRecognition	0.14	0.29	0.25	0.16	0.84	2.4
## FigureRecognition	0.10	0.56	0.22	0.37	0.63	1.4
## ObjectNumber	0.16	0.25	0.39	0.24	0.76	2.1
## NumberFigure	0.04	0.42	0.47	0.39	0.61	2.0
## FigureWord	0.16	0.34	0.28	0.22	0.78	2.4
## Deduction	0.39	0.48	0.18	0.42	0.58	2.2
## NumericalPuzzles	0.19	0.40	0.46	0.41	0.59	2.3
## ProblemReasoning	0.38	0.46	0.21	0.40	0.60	2.4
## SeriesCompletion	0.39	0.53	0.27	0.50	0.50	2.4
## ArithmeticProblems	0.37	0.22	0.55	0.49	0.51	2.1

```
##
##
```

	MR1	MR3	MR2
## SS loadings	3.76	3.37	3.34
## Proportion Var	0.16	0.14	0.14
## Cumulative Var	0.16	0.30	0.44
## Proportion Explained	0.36	0.32	0.32
## Cumulative Proportion	0.36	0.68	1.00

```
##
## Mean item complexity = 1.8
## Test of the hypothesis that 3 factors are sufficient.
##
## The degrees of freedom for the null model are 276 and the objective function was 11.44
## The degrees of freedom for the model are 207 and the objective function was 2.24
##
## The root mean square of the residuals (RMSR) is 0.05
## The df corrected root mean square of the residuals is 0.06
##
## Fit based upon off diagonal values = 0.97
```



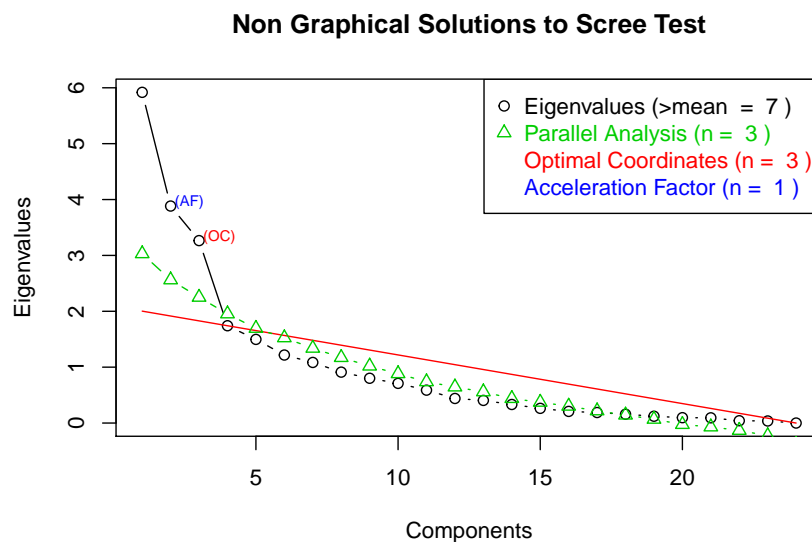
```
## Measures of factor score adequacy
##
## Correlation of (regression) scores with factors    MR1  MR3  MR2
## Multiple R square of scores with factors          0.93 0.88 0.92
## Minimum correlation of possible factor scores      0.74 0.56 0.69
```

mydata pode ser a matriz de dados brutos ou a matriz de covariâncias. Exclusão pareada (*Pairwise deletion*) dos dados faltantes é utilizado e o tipo de rotação pode ser *varimax* ou *promax*.

Determinando o número de Fatores para extrair

Uma decisão crucial na análise fatorial exploratória refere-se a quantos fatores eu devo reter. O pacote **nFactors** oferece um conjunto de funções para auxiliar nesta decisão. Vejamos o *Scree plot*.

```
# Determine Number of Factors to Extract
library(nFactors)
ev <- eigen(cor(mydata)) # get eigenvalues
ap <- parallel(subject=nrow(mydata), var=ncol(mydata),
  rep=100, cent=.05)
nS <- nScree(x=ev$values, aparallel=ap$eigen$qevpea)
plotnScree(nS)
```



Testes usuais de adequação da amostra

Os testes mais usuais para a análise fatorial exploratória são o teste de esfericidade de Bartlett, O teste de K-M-O e a medida de consistência interna Alpha de Crombach. Os três procedimentos são apresentados a seguir e estão disponíveis no pacote **psych** do R.

Bartlett's Test

O teste de Bartlett verifica que a matriz de correlação não é uma matriz identidade. Uma vez que deseja-se criar variáveis latentes (fatores) a partir da combinação de variáveis que meçam o mesmo constructo, é de se esperar que estas sejam correlacionadas. Caso isto ocorra, a matriz de correlações não será uma matriz

identidade (completa de zeros fora da diagonal principal). O teste verifica os resíduos da diferença entre a matriz de correlação e uma matriz identidade.

Como entrada deve ser informado a matriz de correlação dos dados e o tamanho da amostra.

```
library(psych)
cortest.bartlett(mydata, n = 145, diag=TRUE)
```

```
## $chisq
## [1] 1545.862
##
## $p.value
## [1] 2.399561e-175
##
## $df
## [1] 276
```

Kaiser, Meyer, Olkin Measure Of Sampling Adequacy

Henry Kaiser (1970) introduziu uma Medida de Adequação de Amostragem (MSA) de matrizes de dados analíticos de fatores. Kaiser e Rice (1974) então o modificaram. Isso é apenas uma função dos elementos quadrados da matriz imagem em comparação com os quadrados das correlações originais. O MSA geral e as estimativas para cada item são encontrados. O índice é conhecido como índice Kaiser-Meyer-Olkin (KMO).

Como entrada deve ser informado a matriz de correlação dos dados.

```
library(psych)
KMO(mydata)
```

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = mydata)
## Overall MSA = 0.88
## MSA for each item =
```

VisualPerception	Cubes	PaperFormBoard
0.90	0.84	0.78
Flags	GeneralInformation	PargraphComprehension
0.85	0.88	0.89
SentenceCompletion	WordClassification	WordMeaning
0.89	0.92	0.88
Addition	Code	CountingDots
0.81	0.85	0.84
StraightCurvedCapitals	WordRecognition	NumberRecognition
0.89	0.85	0.88
FigureRecognition	ObjectNumber	NumberFigure
0.89	0.85	0.88
FigureWord	Deduction	NumericalPuzzles
0.83	0.93	0.91
ProblemReasoning	SeriesCompletion	ArithmeticProblems
0.93	0.91	0.92

Crombach's Alpha

Alpha é uma das várias estimativas da confiabilidade da consistência interna de um teste. Etenda aqui como teste um questionário ou instrumento de coleta de dados. Como resultado, o teste informa um escore que

aponta a consistência interna das respostas obtidas por meio do instrumento de coleta (o teste). Deve-se imputar a base de dados brutos ou a matriz de correlação em conjunto com o tamanho da amostra.

```
library(psych)
alpha(mydata, n.obs = 145)

##
## Reliability analysis
## Call: alpha(x = mydata, n.obs = 145)
##
##      raw_alpha std.alpha G6(smc) average_r S/N   ase median_r
##          0.91      0.91      0.94        0.3  10 0.011      0.29
##
## lower alpha upper      95% confidence boundaries
## 0.89 0.91 0.93
##
## Reliability if an item is dropped:
##
##      raw_alpha std.alpha G6(smc) average_r S/N alpha se
## VisualPerception      0.91      0.91      0.93      0.30 9.8 0.011
## Cubes                  0.91      0.91      0.94      0.31 10.3 0.011
## PaperFormBoard        0.91      0.91      0.93      0.31 10.2 0.011
## Flags                  0.91      0.91      0.93      0.30 10.1 0.011
## GeneralInformation     0.91      0.91      0.93      0.30 9.7 0.011
## PargraphComprehension  0.91      0.91      0.93      0.30 9.7 0.011
## SentenceCompletion     0.91      0.91      0.93      0.30 9.7 0.011
## WordClassification     0.91      0.91      0.93      0.30 9.7 0.011
## WordMeaning            0.91      0.91      0.93      0.30 9.7 0.011
## Addition               0.91      0.91      0.93      0.31 10.2 0.011
## Code                   0.91      0.91      0.93      0.30 9.9 0.011
## CountingDots           0.91      0.91      0.93      0.31 10.1 0.011
## StraightCurvedCapitals 0.91      0.91      0.93      0.30 9.8 0.011
## WordRecognition        0.91      0.91      0.94      0.31 10.2 0.011
## NumberRecognition      0.91      0.91      0.94      0.31 10.3 0.011
## FigureRecognition      0.91      0.91      0.93      0.30 10.0 0.011
## ObjectNumber           0.91      0.91      0.93      0.30 10.1 0.011
## NumberFigure           0.91      0.91      0.93      0.30 9.9 0.011
## FigureWord             0.91      0.91      0.93      0.31 10.1 0.011
## Deduction              0.91      0.91      0.93      0.30 9.8 0.011
## NumericalPuzzles       0.91      0.91      0.93      0.30 9.8 0.011
## ProblemReasoning       0.91      0.91      0.93      0.30 9.8 0.011
## SeriesCompletion       0.91      0.91      0.93      0.29 9.6 0.011
## ArithmeticProblems     0.91      0.91      0.93      0.30 9.7 0.011
##
##      var.r med.r
## VisualPerception 0.016 0.28
## Cubes             0.015 0.30
## PaperFormBoard    0.015 0.30
## Flags             0.016 0.29
## GeneralInformation 0.014 0.29
## PargraphComprehension 0.014 0.29
## SentenceCompletion 0.014 0.29
## WordClassification 0.015 0.29
## WordMeaning       0.014 0.30
## Addition           0.014 0.30
## Code               0.016 0.29
```

```

## CountingDots          0.015  0.30
## StraightCurvedCapitals 0.016  0.29
## WordRecognition       0.016  0.30
## NumberRecognition     0.016  0.30
## FigureRecognition     0.016  0.29
## ObjectNumber          0.016  0.30
## NumberFigure          0.016  0.29
## FigureWord            0.016  0.30
## Deduction             0.016  0.29
## NumericalPuzzles      0.016  0.28
## ProblemReasoning      0.016  0.29
## SeriesCompletion       0.016  0.28
## ArithmeticProblems    0.016  0.28
##
## Item statistics
##
##           r r.cor r.drop
## VisualPerception    0.62  0.60  0.57
## Cubes               0.42  0.38  0.35
## PaperFormBoard      0.46  0.43  0.40
## Flags              0.51  0.48  0.45
## GeneralInformation  0.67  0.67  0.63
## ParagraphComprehension 0.67  0.66  0.62
## SentenceCompletion  0.65  0.64  0.60
## WordClassification  0.67  0.66  0.63
## WordMeaning         0.66  0.66  0.62
## Addition            0.48  0.47  0.42
## Code               0.59  0.57  0.54
## CountingDots        0.50  0.48  0.44
## StraightCurvedCapitals 0.62  0.61  0.57
## WordRecognition     0.47  0.43  0.41
## NumberRecognition   0.44  0.40  0.38
## FigureRecognition   0.55  0.53  0.50
## ObjectNumber        0.51  0.48  0.45
## NumberFigure        0.57  0.54  0.51
## FigureWord          0.49  0.46  0.43
## Deduction           0.63  0.62  0.59
## NumericalPuzzles    0.62  0.61  0.58
## ProblemReasoning    0.63  0.61  0.58
## SeriesCompletion     0.70  0.69  0.66
## ArithmeticProblems  0.67  0.66  0.63

```

Atividade

A base de dados do MAPEM (arquivo salvo no formato do SPSS) foi obtida por meio de um projeto da PETROBRÁS que buscava avaliar o impacto ambiental da perfuração de poços de petróleo em alto mar. As coletas das amostras de solo ocorreram em períodos de tempo, afastadas por uma dada distância em certa direção a partir do ponto onde um poço de petróleo estava sendo perfurado. São observadas variáveis de identificação do local de coleta e contagem de animais da macro e micro fauna, composição do solo, químicos e sedimentos. Explore a base de dados para identificar os elementos citados.

1. Faça um esboço da localização dos pontos de coleta a partir das informações da base (tenha como origem do sistema o ponto de perfuração do poço).

2. Verifique a adequação do procedimento de análise fatorial utilizando os três testes apresentados no final deste módulo de estudo.
3. Proceda uma análise fatorial exploratória. Apresente os resultados da análise, identificando o perfil de cada fator retido. Descreva sumariamente os procedimentos e decisões realizados na análise.
4. Ao final do arquivo, apresente o código utilizado para a realização de toda a análise.