

# Contents

<i>List of figures</i>	xiii
<i>Preface</i>	xvii
<i>Acknowledgments</i>	xix

## PART I

<b>Multivariate analysis of factors and components</b>	<b>1</b>
<b>1 Factor analysis: Research questions it addresses</b>	<b>3</b>
<i>Introduction</i>	3
<i>Purposes of factor analysis</i>	4
<i>Limitations of factor analysis</i>	5
<i>Common research questions associated with factor analysis</i>	6
<b>2 Assumptions and limitations of factor analysis</b>	<b>13</b>
<i>Introduction</i>	13
<i>Existence of underlying dimensions</i>	13
<i>Proper specification/no selection bias</i>	16
<i>Proper specification of the number of factors</i>	16
<i>Data homogenous on factor structure</i>	17
<i>Valid imputation of factor labels</i>	17
<i>Data level</i>	18
<i>Linearity</i>	22
<i>Multivariate normality</i>	22
<i>Skew and kurtosis</i>	23
<i>Homoskedasticity</i>	25
<i>No influential outliers</i>	25
<i>No influential missing data</i>	30
<i>Moderate to moderate-high intercorrelations without multicollinearity</i>	31
<i>Absence of high multicollinearity</i>	31
<i>No perfect multicollinearity</i>	32
<i>Sphericity</i>	32
<i>Adequate sample size</i>	32

- 3 Fundamental concepts in factor analysis** 34
- Research modes: EFA vs. CFA* 34
  - Estimation methods* 35
  - Data extraction: PCA vs. PFA* 35
  - Other data extraction methods* 38
  - Number of dimensions to extract* 40
  - Item complexity and simple factor structure* 48
  - Rotation of axes* 50
  - Eigenvalues* 55
  - Eigenvectors and factor loadings* 56
  - Communality and uniqueness* 58
  - Factor and component scores* 61
  - Model fit* 63
- 4 Quick start: Principal factor analysis (PFA) in R** 66
- Introduction* 66
  - Setup and example data* 67
  - Parallel analysis* 68
  - Orthogonal PFA with `fa()`* 70
  - Factor scores in `fa()`* 72
  - Beta weights in `fa()`* 73
- 5 Quick start: Confirmatory factor analysis in R** 76
- Overview* 76
  - Testing error in the CFA measurement model* 77
  - Other CFA tests* 77
  - Goodness-of-fit measures* 78
  - Modification indices and parameter change coefficients* 81
  - Path significance and critical ratios* 83
  - R packages for CFA and SEM* 83
  - Example data* 84
  - Creating the CFA model with `lavaan`* 84
  - Residual analysis* 92
  - Modification indices* 93
  - Goodness-of-fit measures* 94
  - Revised model* 96
  - Visualization* 98
- 6 Quick start: Principal component analysis (PCA) in R** 100
- Introduction* 100
  - Setup and data for principal components analysis with `PCA()`* 100
  - Testing factor adequacy with `KMO()`* 102
  - Bartlett's test for sphericity* 102
  - Determining the number of factors to request* 103

Creating the model with <i>PCA()</i>	103
Eigenvalues and the scree plot	104
Empirical scree tests	105
Exploratory graph analysis of factor memberships	107
PCA variable plot	108
Eigenvectors	110
Component loadings	113
Component rotation	114
Biplots and outliers	116
Variable contributions	116
Residual analysis	121
Saving component scores	123
Automatic PCA reporting with “FactoInvestigate”	126
Principal Component Analysis	126
<b>7 Oblique and higher-order factor models</b>	<b>134</b>
Oblique PFA with <i>fa()</i>	134
Oblique PCA with <i>principal()</i>	146
Second-order oblique factor analysis	153
Bifactor models	160
<b>8 Factor analysis for binary, ordinal, and mixed data</b>	<b>165</b>
Polychoric PCA and PFA	165
Heterogeneous PCA with <i>hetcor()</i>	173
Mixed data PCA with <i>PCAmix()</i>	182
Mixed data PCA with <i>FAMD()</i>	208
Mixed data with generalized low-rank models (GLRM)	217
Categorical PCA with <i>prinicals()</i>	238
PCA for binary variables with <i>logisticPCA()</i>	246
<b>9 PFA in greater detail</b>	<b>267</b>
Introduction	267
Extension variables with <i>fa()</i>	267
Orthogonal PFA with <i>factanal()</i>	273
Oblique factor analysis with <i>fa.promax()</i>	282
Bayesian factor analysis with <i>BayesFM</i>	286
Regularized factor analysis with <i>fareg()</i>	292
<b>10 PCA in greater detail</b>	<b>297</b>
Introduction	297
PCA with <i>prcomp()</i>	297
PCA with <i>principal()</i>	337
PCA for R with <i>princomp()</i>	351

## PART II

<b>Additional tools for dimension reduction</b>	<b>355</b>
<b>11 Sixteen additional methods for dimension reduction (dimRed)</b>	<b>357</b>
<i>Dimension reduction in the dimRed() system</i>	<i>357</i>
<i>Introduction</i>	<i>357</i>
<i>Setup</i>	<i>357</i>
<i>The embed() function in dimRed</i>	<i>358</i>
<i>Dimension reduction methods in dimRed</i>	<i>359</i>
PCA and PCA_L1 methods	360
Kernel PCA (kPCA)	362
Classical multidimensional scaling	365
Non-metric multidimensional scaling	367
Locally linear embedding	369
Hessian locally linear embedding	372
Laplacian eigenmaps	373
Diffusion maps	376
Force directed methods	379
t-Distributed stochastic neighbor embedding (tSNE)	384
Independent component analysis (FastICA)	385
Dimensionality reduction via regression (DRR)	387
Non-negative matrix factorization (NNMF)	391
Isomap	393
Autoencoder	396
<b>12 Metrics for comparing and evaluating dimension reduction models</b>	<b>399</b>
<i>Performance quality metrics for dimRed</i>	<i>399</i>
<i>Multi-method multi-measure comparison</i>	<i>406</i>
<i>The “coRanking” package</i>	<i>408</i>
<i>Package dimRED multi-method multi-measure comparison with custom parameters</i>	<i>414</i>
<b>13 Recipes: An alternative system for dimension reduction</b>	<b>419</b>
<i>Introduction</i>	<i>419</i>
<i>The recipes design framework</i>	<i>419</i>
<i>Libraries and setup for this section</i>	<i>420</i>
<i>The unvotes example data</i>	<i>421</i>
<i>Data levels: a note of caution</i>	<i>422</i>
<i>Illustration of use of the unvotes data</i>	<i>423</i>
<i>PCA: standard deviations, variances, eigenvectors, eigenvalues, contributions, and loadings</i>	<i>425</i>
<i>PCA with the recipes package</i>	<i>435</i>
<i>ICA with the recipes package</i>	<i>449</i>
<i>KPCA with the recipes package</i>	<i>454</i>

<b>14 Factor analysis for neural models</b>	466
Introduction	466
Example data and setup	466
PCA in caret pre-processing	467
Use of PCA in the pcaNNET modeling method	472
Autoencoder with dimRed	478
<b>15 Factor analysis for time series data</b>	479
Introduction	479
Setup	480
Example data	481
Visualizing longitudinal data with the “ggplot2” package	481
Visualizing longitudinal data with the “brolgar” package	482
Data preparation for FPCA for longitudinal data	483
Number of components	486
Component trends over time	487
Diagnostic plots	487
Outlier detection	489
Scores	491
Other R packages for functional PCA	492
<b>Appendix 1: Datasets used in this volume</b>	493
<b>Appendix 2: Introduction to R and RStudio</b>	497
Why R?	497
Installing R and RStudio	498
Example data	498
Quick start: computing a correlation	500
Importing data	502
Saving data	506
Adding value labels to data	507
Inspecting data	509
R data structures	514
Handling missing values	519
Finding useful packages to install	523
Installing packages	525
Updating packages	528
Using, saving, and loading packages and sessions	529
Visualization and graphics in R	531
Data management basics	532
Dealing with error messages	534
Obtaining data	536

Getting help 537

A note on using the `attach()` command 539

**Appendix 3: Frequently asked questions** 540

How to report factor analysis 540

What are “data modes” in factor analysis? 540

What is KMO? What is it used for? 541

Is it necessary to standardize one’s variables before applying factor analysis? 542

Can you pool data from two samples together in factor analysis? 542

How does factor comparison of the factor structure of two samples work? 542

**References** 545

**Index** 555