ISSN NO. 2320-5407

*Journal homepage: http://www.journalijar.com*

*INTERNATIONAL JOURNAL OF ADVANCED RESEARCH*

**RESEARCH ARTICLE**

# COMPARATIVE STUDY OF PRINCIPAL COMPONENTS AND FACTOR ANALYTIC TECHNIQUES

**REUBEN Benham Zangaluka and TORSEN Emmanuel**

Department of Statistics and Operations Research, Modibbo Adama University of Technology, Yola, Adamawa-Nigeria.

## *Manuscript Info*

***Manuscript History:***

Received: 25 April 2015
Final Accepted: 14 May 2015
Published Online: June 2015

***Key words:***

Principal Components, Factor analytic techniques, Variance, Correlation

***Corresponding Author***

**REUBEN, Benham Zangaluka**

## *Abstract*

Principal Components and factor analytic techniques take large number of variables and reduce them to much smaller number of coherent subsets such that variables within a subset are related to one another but independent to those in other subsets. These methods summarize the pattern of correlation between observed variables. In this paper, principal components and factor analytic techniques are compared using data from Nigerian Consumption Pattern 2009/2010. The results revealed that factor analytic techniques preserve correlation more than principal components, while on the other hand, principal components preserve variance more than factor analytic techniques. We therefore conclude that factor analysis should be used when interest is placed on making statements about the factors that are responsible for a set of observed responses, and principal component analysis should be used when interest is based on performing data reduction.



## INTRODUCTION

Factor Analysis (FA) is often confused with Principal Component Analysis (PCA), a similar statistical procedure. However, there are significant differences between the two: factor analysis and principal component analysis will provide somewhat different results when applied to the same data. The purpose of PCA is to derive a relatively small number of components that can account for the variability found in a relatively large number of measures. This procedure, called data reduction, is typically performed when a researcher does not want to include all the original measures in analysis but still wants to work with the information that they contain (DeCoster, 1998). On the other hand, the primary objectives of factor analysis are to determine the number of common factors influencing a set of measures, and the strength of the relationship between each factor and each observed measure. Because it is a variable reduction procedure, principal component analysis is similar in many respects to exploratory factor analysis. In fact, the steps followed when conducting a PCA, are virtually identical to those followed when conducting an exploratory factor analytical techniques. However, there are significant conceptual differences between the two procedures, and it is important that we do not mistakenly claim that we are performing factor analysis when actually performing principal component analysis. The most important conceptual difference between the procedures deals with the assumption of an underlying causal structure: factor analysis assumes that the co-variation in the observed variables is due to the presence of one or more latent variables (factors) that exert causal influence on those observed variables (Kline,1994). In contrast, PCA makes no assumption about the underlying causal model. PCA is simply a variable reduction procedure that (typically) results in a relatively small number of components that account for most of the variance in a set of observed variables.

Factor analysis, like principal component analysis, attempts to explain a set of data in terms of a smaller number of dimensions that one begins with, but the procedures used to achieve this goal are essentially different in the two methods. Factor analysis unlike principal component analysis, begins with a hypothesis about the covariance (or correlational) structure of the variables (Landau and Everitt, 2004).

Timm (2002) concluded that the biggest difference between principal components and factor analysis comes from model philosophy. Factor analysis imposes a strict structure of a fixed number of common (latent) factors whereas the principal component analysis determines a given number of components in decreasing order of importance.

Simplistically, though, factor analysis derives a mathematical model from which factors are estimated, whereas PCA merely decomposes the original data into a set of linear variate (Dunteman, 1989). Guadagnoli and Velicer (1988) concluded that the solutions generated from principal component analysis differ little from those derived from factor analytic techniques.

Brown (2009) shows what PCA and FA are, and in part, how they should be presented and interpreted. In the process, he had defined and exemplified loadings, communalities, proportion of variance, components, factors, PCA and FA. He also went further to explore the basic mathematical and conceptual differences between PCA and FA, and discussed how researchers decide on whether to use PCA or FA.

From an implementation point of view, the PCA is based on a well-defined, unique algorithm (spectral decomposition), whereas fitting a factor analysis model involves a variety of analysis procedure which opens the door for subjective interpretation and yields therefore a spectrum of results. This data analysis philosophy makes factor analysis difficult especially if the model specification involves cross-validation and a data-driven selection of the number of factors (Simar and Hardle, 2007). PCA solved a problem similar to the problem of common factor analysis, but different enough to lead to confusion (Richard, 2004).

Though there are some important conceptual differences between PCA and FA that have been investigated by a number of researchers, this paper attempts to accentuate some of these differences by highlighting how each tool behaves at each stage of computation.

## 2.0 MATERIALS AND METHODOLOGY

### 2.1 SOURCE OF DATA

The data used in this research were secondary data on non-food commodity expenditure for 36 states of Nigeria and the Federal Capital Territory (FCT), Abuja. The data were sourced from the National Bureau of Statistics, Preliminary Report of Consumption Pattern in Nigeria for the year 2009/2010.

The variables are defined as follows: $X_1$= Clothing and footwear; $X_2$ = Rent; $X_3$ = Fuel/Light; $X_4$ = Household Goods; $X_5$ = Health Expenditure; $X_6$ = Transport; $X_7$ = Education Expenditure; $X_8$= Entertainment; $X_9$ = Water; $X_{10}$ = other services

### 2.2 DATA ANALYSIS

Principal component analysis and factor analysis were conducted on the data using SPSS 16.0. The variables were measured on the same experimental unit of percentages of non-food commodity expenditure of 36 states of the Federal Republic of Nigeria and the Federal Capital Territory (FCT), Abuja. Correlations between variables were obtained. Kaiser criterion was used as the method for determining the optimal number of factors or components for inclusion. Principal components extraction and Maximum likelihood extraction were used to extract components and factors respectively. Varimax orthogonal rotation which produces uncorrelated components/factors was used to rotate factors or components to obtain final solution that aid interpretation.

After obtaining the output of both analyses, the solutions were used to compare the computational efficiency of principal components analysis and factor analysis based on the following criteria:

**(i) Data Fitness:** Residuals of the model are the differences between the matrix based on the model and matrix based on observed data. SPSS 16.0 produces these residuals in the lower table of the reproduced matrix and it is expected to be relatively few of these values to be greater than 5% (Field, 2004). The higher the residuals the less fit a dataset is to the model.

**(ii) Variance Maximization:** The eigenvalue associated with each component (factor) represent the variance explained by that particular component/factor. A model is consistent if the cumulative value of the retained components (factors) is the same before and after rotation (Field, 2004). Rotation has the effect of optimizing the factor structure and its consequence is that the relative importance of the retained factors (components) is equalized.

Test of normality was also conducted on the data and it was found that most variables are normally distributed as their significances are greater than 5% (see Appendix I).

## 3.0 RESULTS AND DISCUSSION

Principal component analysis and factor analysis (using maximum likelihood factor extraction) display the same result for descriptive statistics as both analyses are performed by examining the patterns of correlations or covariation between the observed measures.

**Table 1: Eigenvalues and Communalities**

| Variables | PCA | | | | FA | | | |
|---|---|---|---|---|---|---|---|---|
| | Comp1 | Comp2 | Comp3 | $h^2$ | Factor1 | Factor2 | Factor3 | $h^2$ |
| $X_1$ | -0.121 | **0.909** | 0.194 | 0.878 | -0.072 | **0.978** | -0.193 | 0.999 |
| $X_2$ | -0.227 | **-0.836** | **0.340** | 0.865 | -0.224 | **-0.821** | **-0.524** | 0.999 |
| $X_3$ | -0.135 | -0.046 | **-0.841** | 0.727 | -0.046 | -0.096 | **0.686** | 0.482 |
| $X_4$ | **-0.916** | 0.152 | -0.161 | 0.889 | **-0.961** | 0.160 | 0.222 | 0.999 |
| $X_5$ | 0.131 | 0.696 | 0.514 | 0.766 | 0.129 | **0.586** | -0.166 | 0.387 |
| $X_6$ | **0.827** | 0.026 | -0.206 | 0.727 | **0.772** | -0.033 | 0.385 | 0.745 |
| $X_7$ | 0.286 | **-0.563** | 0.359 | 0.528 | 0.221 | -0.435 | -0.137 | 0.257 |
| $X_8$ | **0.543** | 0.333 | -0.420 | 0.586 | **0.526** | 0.229 | 0.210 | 0.373 |
| $X_9$ | -0.285 | -0.057 | **0.656** | 0.515 | -0.159 | -0.037 | **-0.487** | 0.264 |
| $X_{10}$ | **0.941** | 0.007 | 0.31 | 0.886 | **0.923** | -0.017 | 0.055 | 0.856 |
| % of variance | 0.3044 | 0.2458 | 0.1866 | **0.7368** | 0.2798 | 0.2253 | 0.1311 | **0.6362** |

Table 1 shows PCA and FA analyses with Varimax rotation, Eigenvalue ≥ 1 and the resulting loadings for the percentage of non-food commodities in Nigeria for the year 2009/2010. Observe that the first column has 10 variables as earlier defined, then the next three columns show the results for a PCA of the data, and the last three columns show corresponding results for an FA of the same data. Observe that the actual loadings differ for the PCA and FA. Note also that the pattern of relatively strong loadings are the same for both analyses, so in that sense, it made little difference which analysis was used. However, loadings, higher communalities, and ultimately the proportion of variance accounted for in the aggregate is 73.68% in PCA as opposed to 63.62% in FA. This is because FA excludes unique variances which are used in the PCA to contribute to higher loadings with the components in ways that are not present in FA.

**Table 2: Total Variance Explained by Retained Components**

| | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| Component | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 3.044 | 30.444 | 30.444 | 3.044 | 30.444 | 30.444 | 2.969 | 29.689 | 29.689 |
| 2 | 2.458 | 24.579 | 55.023 | 2.458 | 24.579 | 55.023 | 2.470 | 24.702 | 54.391 |
| 3 | 1.866 | 18.656 | 73.679 | 1.866 | 18.656 | 73.679 | 1.929 | 19.288 | 73.679 |
| 4 | .753 | 7.526 | 81.204 | | | | | | |
| 5 | .637 | 6.372 | 87.576 | | | | | | |
| 6 | .545 | 5.450 | 93.025 | | | | | | |
| 7 | .336 | 3.357 | 96.382 | | | | | | |
| 8 | .250 | 2.500 | 98.882 | | | | | | |
| 9 | .111 | 1.112 | 99.994 | | | | | | |
| 10 | .001 | .006 | 100.000 | | | | | | |

Extraction Method: Principal Component Analysis.

**Table 3: Total Variance Explained by Retained Factors**

| | Initial Eigenvalues | | | Extraction Sums of Squared Loadings | | | Rotation Sums of Squared Loadings | | |
|---|---|---|---|---|---|---|---|---|---|
| Factor | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % | Total | % of Variance | Cumulative % |
| 1 | 3.044 | 30.444 | 30.444 | 2.291 | 22.914 | 22.914 | 2.798 | 27.975 | 27.975 |
| 2 | 2.458 | 24.579 | 55.023 | 2.722 | 27.216 | 50.130 | 2.253 | 22.534 | 50.509 |
| 3 | 1.866 | 18.656 | 73.679 | 1.349 | 13.487 | 63.617 | 1.311 | 13.108 | 63.617 |
| 4 | .753 | 7.526 | 81.204 | | | | | | |
| 5 | .637 | 6.372 | 87.576 | | | | | | |
| 6 | .545 | 5.450 | 93.025 | | | | | | |
| 7 | .336 | 3.357 | 96.382 | | | | | | |
| 8 | .250 | 2.500 | 98.882 | | | | | | |
| 9 | .111 | 1.112 | 99.994 | | | | | | |
| 10 | .001 | .006 | 100.000 | | | | | | |

Extraction Method: Maximum Likelihood.

Table 2 and Table 3 show the importance of the ten principal components (factors). Only the first three have eigenvalues over 1.00, and together these explain over 73% of the total variability in the data. While this figure remains the same after extraction using PCA, it differs by about 10% after rotation using FA (maximum likelihood) extraction.

Appendix II and Appendix III shows the reproduced correlation of PCA and FA respectively. The reproduced correlations matrix differs from those in the observed matrix because they stem from the model. Therefore, to assess the fit of the model we can look at the differences between the observed correlations and the correlations based on the model. The difference can be calculated as follows:

Residual = observed correlation – correlation from model.

Note that this difference is the value quoted in the lower half of the reproduced matrix (labeled residuals). Therefore, the lower half of the reproduced matrix contains the differences between the observed correlation coefficients and the ones predicted from the model. For a good model these values will all be small. In fact, most values should be less than 0.05 significant levels (Field, 2004). Rather than scan this huge matrix, SPSS 16.0 provides a footnote summary, which states how many residuals have absolute value greater than 0.05. For these data values there are 14 (31%) absolute values greater than 0.05 in FA and 23(51%) absolute values greater than 0.05 in PCA. There are no hard and fast rules about what proportion of residuals should be below 0.05; however, if more than 50% are greater than 0.05 we probably have grounds for concern (Field, 2004).

**APPENDICES**

**Appendix I**: **Tests of Normality**

| | Kolmogorov-Smirnov[a] | | | Shapiro-Wilk | | |
|---|---|---|---|---|---|---|
| Var. | Statistic | Df | Sig. | Statistic | Df | Sig. |
| X1 | .095 | 33 | .200* | .977 | 33 | .684 |
| X2 | .116 | 33 | .200* | .935 | 33 | .049 |
| X3 | .100 | 33 | .200* | .957 | 33 | .209 |
| X4 | .145 | 33 | .074 | .945 | 33 | .097 |
| X5 | .170 | 33 | .016 | .953 | 33 | .167 |
| X6 | .100 | 33 | .200* | .967 | 33 | .406 |
| X7 | .113 | 33 | .200* | .940 | 33 | .069 |
| X8 | .165 | 33 | .024 | .754 | 33 | .000 |
| X9 | .119 | 33 | .200* | .904 | 33 | .007 |
| X10 | .163 | 33 | .026 | .939 | 33 | .065 |

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

There are a number of different extraction methods in factor analysis, unless there is a serious lack of multivariate normality, maximum likelihood extraction is the best among these method (DeCoster, 1998), hence the need to normality test.

The Kolmogorov-Smirnov and Shapiro-Wilk tests compare the scores in the sample to a normally distributed set of scores with the same mean and standard deviation. If the test is non-significant ($p > 0.05$) it tells us that the distribution of the sample is not significantly different from a normal distribution (Field, 2004).

The Appendix I includes the test statistic itself, the degrees of freedom and the significance value of the test. A significant value less than 0.05 indicate a deviation from normality. Both tests are highly non-significant, indicating that the distribution is normal.

**Appendix II: Reproduced Correlations (PCA)**

| Var | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 |
|---|---|---|---|---|---|---|---|---|---|---|
| X1 | .878[a] | -.666 | -.189 | .217 | .717 | -.117 | -.477 | .160 | .110 | -.101 |
| X2 | -.666 | .865[a] | -.217 | .027 | -.437 | -.279 | .528 | -.549 | .335 | -.209 |
| X3 | -.189 | -.217 | .727[a] | .252 | -.482 | .060 | -.314 | .264 | -.511 | -.153 |
| X4 | .217 | .027 | .252 | .889[a] | -.097 | -.721 | -.406 | -.379 | .147 | -.866 |
| X5 | .717 | -.437 | -.482 | -.097 | .766[a] | .020 | -.171 | .091 | .261 | .144 |
| X6 | -.117 | -.279 | .060 | -.721 | .020 | .727[a] | .148 | .544 | -.372 | .772 |
| X7 | -.477 | .528 | -.314 | -.406 | -.171 | .148 | .528[a] | -.186 | .186 | .277 |
| X8 | .160 | -.549 | .264 | -.379 | .091 | .544 | -.186 | .586[a] | -.449 | .501 |
| X9 | .110 | .335 | -.511 | .147 | .261 | -.372 | .186 | -.449 | .515[a] | -.248 |
| X10 | -.101 | -.209 | -.153 | -.866 | .144 | .772 | .277 | .501 | -.248 | .886[a] |
| X1 | | -.020 | -.035 | -.034 | -.122 | -.047 | .061 | -.014 | -.041 | .008 |
| X2 | -.020 | | -.055 | -.059 | .013 | -.068 | -.149 | .132 | -.014 | -.013 |
| X3 | -.035 | -.055 | | -.072 | .104 | -.042 | .093 | .008 | .230 | .010 |
| X4 | -.034 | -.059 | -.072 | | .030 | .059 | .093 | -.044 | -.108 | -.012 |
| X5 | -.122 | .013 | .104 | .030 | | -.018 | .090 | .024 | -.051 | -.041 |
| X6 | -.047 | -.068 | -.042 | .059 | -.018 | | -.080 | -.184 | .040 | -.067 |
| X7 | .061 | -.149 | .093 | .093 | .090 | -.080 | | .025 | -.134 | -.065 |
| X8 | -.014 | .132 | .008 | -.044 | .024 | -.184 | .025 | | .114 | -.066 |
| X9 | -.041 | -.014 | .230 | -.108 | -.051 | .040 | -.134 | .114 | | .013 |
| X10 | .008 | -.013 | .010 | -.012 | -.041 | -.067 | -.065 | -.066 | .013 | |

Extraction Method: Principal Component Analysis.

a. Reproduced communalities

b. Residuals are computed between observed and reproduced correlations. There are 23 (51.0%) non redundant residuals with absolute values greater than 0.05.

**Appendix III: Reproduced Correlations (FA)**

| Var | X1 | X2 | X3 | X4 | X5 | X6 | X7 | X8 | X9 | X10 |
|---|---|---|---|---|---|---|---|---|---|---|
| X1 | .999[a] | -.686 | -.223 | .183 | .596 | -.163 | -.415 | .146 | .070 | -.093 |
| X2 | -.686 | .999[a] | -.271 | -.033 | -.423 | -.347 | .380 | -.416 | .321 | -.222 |
| X3 | -.223 | -.271 | .482[a] | .181 | -.176 | .232 | -.063 | .098 | -.323 | -.003 |
| X4 | .183 | -.033 | .181 | .999[a] | -.067 | -.662 | -.313 | -.422 | .039 | -.878 |
| X5 | .596 | -.423 | -.176 | -.067 | .387[a] | .016 | -.203 | .167 | .039 | .100 |
| X6 | -.163 | -.347 | .232 | -.662 | .016 | .745[a] | .133 | .479 | -.309 | .734 |
| X7 | -.415 | .380 | -.063 | -.313 | -.203 | .133 | .257[a] | -.012 | .047 | .204 |
| X8 | .146 | -.416 | .098 | -.422 | .167 | .479 | -.012 | .373[a] | -.195 | .493 |
| X9 | .070 | .321 | -.323 | .039 | .039 | -.309 | .047 | -.195 | .264[a] | -.173 |
| X10 | -.093 | -.222 | -.003 | -.878 | .100 | .734 | .204 | .493 | -.173 | .856[a] |
| X1 | | -2.325E-6 | .000 | -2.538E-6 | .000 | .000 | .000 | 6.052E-5 | .000 | .000 |
| X2 | -2.325E-6 | | .000 | -2.873E-6 | .000 | .000 | .000 | -4.266E-5 | .000 | .000 |
| X3 | .000 | .000 | | -.001 | -.201 | -.213 | -.159 | .174 | .043 | -.140 |
| X4 | -2.538E-6 | -2.873E-6 | -.001 | | .000 | .000 | -7.064E-5 | .000 | .000 | -1.997E-5 |
| X5 | .000 | .000 | -.201 | .000 | | -.014 | .123 | -.053 | .171 | .002 |
| X6 | .000 | .000 | -.213 | .000 | -.014 | | -.064 | -.118 | -.023 | -.029 |
| X7 | .000 | .000 | -.159 | -7.064E-5 | .123 | -.064 | | -.148 | .004 | .007 |
| X8 | 6.052E-5 | -4.266E-5 | .174 | .000 | -.053 | -.118 | -.148 | | -.141 | -.058 |
| X9 | .000 | .000 | .043 | .000 | .171 | -.023 | .004 | -.141 | | -.061 |
| X10 | .000 | .000 | -.140 | -1.997E-5 | .002 | -.029 | .007 | -.058 | -.061 | |

Extraction Method: Maximum Likelihood.

b. Residuals are computed between observed and reproduced correlations. There are 14 (31.0%) non redundant residuals with absolute values greater than 0.05.

a. Reproduced communalities

**Appendix IV: Correlation Coefficients**

Correlation Matrix

| Variables | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $X_8$ | $X_9$ | $X_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $X_1$ | 1.000 | -.668 | -.181 | .325 | .579 | -.336 | -.371 | .108 | .074 | -.247 |
| $X_2$ | -.668 | 1.000 | -.339 | -.153 | -.360 | -.245 | .326 | -.449 | .352 | -.100 |
| $X_3$ | -.181 | -.339 | 1.000 | .052 | -.347 | .147 | -.245 | .406 | -.311 | -.039 |
| $X_4$ | .325 | -.153 | .052 | 1.000 | .104 | -.608 | -.393 | -.398 | .043 | -.872 |
| $X_5$ | .579 | -.360 | -.347 | .104 | 1.000 | -.228 | -.038 | .074 | .232 | -.069 |
| $X_6$ | -.336 | -.245 | .147 | -.608 | -.228 | 1.000 | .163 | .350 | -.372 | .661 |
| $X_7$ | -.371 | .326 | -.245 | -.393 | -.038 | .163 | 1.000 | -.160 | .055 | .297 |
| $X_8$ | .108 | -.449 | .406 | -.398 | .074 | .350 | -.160 | 1.000 | -.343 | .435 |
| $X_9$ | .074 | .352 | -.311 | .043 | .232 | -.372 | .055 | -.343 | 1.000 | -.262 |
| $X_{10}$ | -.247 | -.100 | -.039 | -.872 | -.069 | .661 | .297 | .435 | -.262 | 1.000 |

The correlation coefficients show the relationship between the observed measures. The correlation matrix of principal components is the same to those of factor analytical techniques. It is from this matrix that reproduced correlations for both techniques are estimated (see Appendix II and Appendix III).

## 4.0 CONCLUSION

Factor Analytical techniques (FA) preserve correlations more than Principal Component Analysis (PCA) with a small residuals of 31% which show that there is a very little difference between the reproduced correlations and

the correlations actually observed between the variables against the residuals of PCA which stands at 51%. PCA preserve more variability of the original data set at 73.68% over FA which preserves variability of the original data set at 63.62%.

While FA preserves more correlations than PCA, the later accounts for more variance in the observed variables than the former. Therefore, when it is needful to do a data reduction, PCA should be used; but when a statement about the underlying causal structure is desirable, FA should be used.

This work has further deepened the discrepancies between PCA and FA, thereby providing assistance to researchers to ease their decision making as to which technique to use apriori.

## REFERENCES

Brown, J.D. (2009): Principal Component Analysis and Exploratory Factor Analysis-Definition, Differences and Choices: JALT Testing and Evaluation. SIG Newsletter. Vol. 13, Number 1, PP 26-30. Retrieved April 16, 2012. Retrieved from Shiken database.

DeCoster, J. (1998). Overview of Factor Analysis. PP 1-3. Retrieved June 14, 2011, from http://www.stat-help.com

Dunteman, G. H. (1989). Principal Component Analysis: Quantitative Applications in the Social Sciences. Newbury Park, California: Sage Publication. PP 7

Field, A. (2004). Discovering Statistics Using SPSS, Second Edition. : Sage Publications. London. PP 641-665

Guadognoli, E. and Velicer, W.F. (1988). Relation of Sample Sizes to the Stability of Component Patterns. Psychological Bulletin. PP 265-275.

Kline, P. (1994). An Easy Guide to Factor Analysis. London: Routledge.

Landau, S. and Everitt, B.S. (2004). Statistical Analysis Using SPSS. Chapman and Hall/CRC Press LLC. London. PP 213.

National Bureau of Statistics (2010). Consumption Pattern in Nigeria (2009/2010). Retrieved March 20, 2010, from www.nigerianstat.gov.ng

Richard, B. D. (2004): Factor Analysis. Retrieved July 2004, from http://www.chass.ncsu.edu/garson

Simar, L. and Hardle, W. (2007). Applied Multivariate Statistical Analysis. Berlin Heidelberg: Springer-Verlag. PP 265

Timm, N.H. (2002). Applied Multivariate Analysis: Springer-Verlag Inc. New York. Page 445.