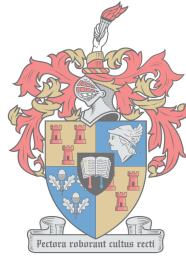


An application of copulas to improve PCA biplots for multivariate extremes

by

Justin Perrang



UNIVERSITEIT
iYUNIVESITHI
STELLENBOSCH
UNIVERSITY

*Research assignment presented in partial fulfilment of the
requirements for the degree of Master of Commerce
(Financial Risk Management) in the Faculty of Economic and
Management Sciences at Stellenbosch University*

Supervisor: Mr. C.J. Van Der Merwe

December 2018

DECLARATION

By submitting this research assignment electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: _____

Justin Perrang

Date: _____

ABSTRACT

Principal Component Analysis (PCA) biplots is a valuable means of visualising high dimensional data. The application of PCA biplots over a wide variety of research areas containing multivariate data is well documented. However, the application of biplots to financial data is limited. This is partly due to PCA being an inadequate means of dimension reduction for multivariate data that is subject to extremes. This implies that its application to financial data is greatly diminished since extreme observations are common in financial data. Hence, the purpose of this research is to develop a method to accommodate PCA biplots for multivariate data containing extreme observations. This is achieved by fitting an elliptical copula to the data and deriving a correlation matrix from the copula parameters. The copula parameters are estimated from only extreme observations and as such the derived correlation matrices contain the dependencies of extreme observations. Finally, applying PCA to such an “extremal” correlation matrix more efficiently preserves the relationships underlying the extremes and a more refined PCA biplot can be constructed.

Opsomming

Hoofkomponent Analise (HKA) bistippings is 'n nuttige metode om meer dimensionele data te visualiseer. Die toepassing van HKA bistippings is al goed gedokumenteer oor 'n wye verskeidenheid van navorsingsareas waar meerveranderlike data voorkom, maar die toepassing van bistippings op finansiële data is beperk. Dit is deels te wyte aan HKA wat 'n onvoldoende metode is van dimensie reduksie van meerveranderlike data wat ekstreme waarnemings bevat. Dit impliseer dat die toepassing daarvan op finansiële data aansienlik beperk is, gegee dat ekstreme waarnemings algemeen voorkom in finansiële data. Die doel van hierdie navorsing is om 'n metode te ontwikkel om HKA- bistippings toe te pas op meerveranderlike data wat ekstreme waarnemings bevat. Dit word gedoen deur 'n elliptiese copula op die data te pas en 'n korrelasiematriks uit die copula parameters af te lei. Die copula parameters word bepaal deur slegs die ekstreme waarnemings te gebruik en dus dui die afgeleide korrelasiematrikse die afhanklikhede van slegs ekstreme waarnemings aan. Laastens, deur HKA op so 'n "ekstreme" korrelasie matriks toe te pas, word die verwantskappe onderliggend aan die ekstreme waardes meer doeltreffend behou en kan 'n meer onderskeidende HKA bistipping gekonstrueer word.

ACKNOWLEDGEMENTS

Firstly, I would like to acknowledge my supervisor, Carel van der Merwe, for his guidance and inspiration throughout my postgraduate studies. Your mentorship over the past few years has been invaluable. It was a pleasure learning from you!

Also, Professor T. De Wet, thank you for the time and effort you invested in guiding me during the year.

And, Professor W.J. Conradie, thank you for introducing me to the wonderful world of Financial Risk Management and for all the guidance you provided during my years at Stellenbosch University.

Furthermore, I would like to thank all the staff in the Department of Statistics and Actuarial Science for all the support during my studies.

I would also like to extend my sincerest gratitude to...

Schroders Asset Management for funding my postgraduate studies.

My Parents, Gabriel and Elsabe, for their caring love, encouragement and support throughout my life.

My Girlfriend, Mighial Adams, for the love, support and, motivation through thick and thin.

My Classmates, Monique, Ryan, Jan, and Nadia, the last two years was a blast.

And Finally to my friends, Clay, Eaden, Jacques, Leonardo and Aldrin for the encouragement and memories.

CONTENTS

DECLARATION	i
ABSTRACT	ii
Opsomming	iii
ACKNOWLEDGEMENTS	iv
CONTENTS	v
LIST OF FIGURES	viii
LIST OF TABLES	x
1 INTRODUCTION	1
1.1 BACKGROUND AND MOTIVATION	1
2 A REVIEW OF PCA BIPLOTS	6
2.1 PRINCIPAL COMPONENT ANALYSIS (PCA)	6
2.2 PCA BILOT CONSTRUCTION	8
2.3 PCA BILOT INTERPRETATION	10
2.4 PCA BILOT QUALITY MEASURES	15
2.5 SUMMARY	17

3	A REVIEW OF COPULAS AND DEPENDENCE	18
3.1	DEPENDENCE MEASURES	19
3.2	COPULA THEORY	22
3.2.1	Elliptical copulas	26
3.2.2	Archimedean copulas	32
3.3	MULTIVARIATE EXTREME DEPENDENCE ANALYSIS	37
3.4	SUMMARY	39
4	METHODOLOGY	40
4.1	THE REFINED PCA BILOT	40
4.2	SIMULATION DESIGN	42
4.2.1	Data generation	43
4.2.2	Traditional and Refined PCA biplot quality measures	44
4.2.3	Testing procedure	45
4.3	SUMMARY	45
5	SIMULATION RESULTS	48
5.1	GAUSSIAN COPULA WITH GAMMA MARGINALS	49
5.2	GUMBEL COPULA WITH GAMMA MARGINALS	55
5.3	GUMBEL COPULA WITH HETEROGENEOUS MARGINALS	61
5.4	SUMMARY	66
6	FINANCIAL APPLICATION	67
6.1	FOREIGN EXCHANGE APPLICATION	67
6.2	SUMMARY	73
7	CONCLUSION AND OPEN QUESTIONS	75

ADDENDA	78
A REQUIRED LINEAR ALGEBRA RESULTS	79
A.1 Spectral decomposition	79
A.2 Singular value decomposition (SVD)	80
B DATA SETS	81
C R-CODE	86
C.1 Code for 95% VaR example	86
C.2 Code to construct Refined biplot adjusted from UBbipl	88
C.3 Code for biplot simulation engine	92
C.4 Code for application of Refined biplots	96
LIST OF REFERENCES	110

LIST OF FIGURES

2.1	Biplot for 95% VaR data.	12
2.2	Correlation biplot for 95% VaR data.	13
2.3	Biplot for 95% VaR data with predictions of day 16 VaR.	14
3.1	Simulation of 10000 bivariate normal and t_4 distributed random variables .	31
3.2	Scatter plot of univariate marginals corresponding to the Gaussian and t_4 - Copula	32
3.3	Scatter plots of univariate marginals corresponding to a Gumbel copula . .	36
3.4	Scatter plots of univariate marginals corresponding to a Clayton copula . .	36
3.5	Scatter plots of univariate marginals corresponding to a Frank copula	37
4.1	Flow diagram of the simulation structure	47
5.1	Pairs scatter plot of a 5-variate Gamma-Gaussian distribution	50
5.2	Traditional (a) and Refined (b) biplots for 5000 observations from a 5-variate Gamma-Gaussian distribution.	51
5.3	Pairs scatter plot of a 5-variate Gamma-Gumbel distribution	56
5.4	Traditional (a) and Refined (b) biplots for 5000 observations from a 5-variate Gamma-Gumbel distribution.	57
5.5	Pairs scatter plot of a 4-variate Heterogeneous-Gumbel distribution	62

5.6	Traditional (a) and Refined (b) biplots for 5000 observations from a 4-variate Heterogeneous-Gumbel distribution.	63
6.1	Pairs scatter plot for Rand/Foreign currency exchange rate monthly returns	68
6.2	Traditional biplot for Rand/Foreign currency exchange rate monthly returns	69
6.3	Refined biplot for Rand/Foreign currency exchange rate monthly returns . .	69
6.4	Traditional biplot prediction Rand/Foreign currency exchange rate monthly return on January 2016	72
6.5	Traditional biplot prediction Rand/Foreign currency exchange rate monthly return on January 2016	73

LIST OF TABLES

2.1	Van Blerk (2000) 95% VaR of financial trading desks	11
2.2	Correlation matrix for 95% VaR of 7 financial trading desks	14
2.3	Actual and predicted value of the 95% VaR for day 16	15
3.1	Popular families of Archimedean copulas	34
5.1	Predictivity and Adequacy measures for biplots constructed from 5000 observations from a 5-variate Gamma-Gaussian distribution.	52
5.2	Overall sample error for biplots constructed from 5000 observations from a 5-variate Gamma-Gaussian distribution.	52
5.3	Extreme sample error for biplots constructed from 5000 observations from a 5-variate Gamma-Gaussian distribution.	53
5.4	Overall and extreme sample error simulation results for biplots from a Gamma-Gaussian distribution. Each table consists of the average sample error, standard error and p-values for 4, 5 and 7 variables using 500 and 5000 observations for 20 and 80 tail samples.	54
5.5	Predictivity and Adequacy measures for biplots constructed from 5000 observations from a 5-variate Gamma-Gumbel distribution.	57
5.6	Overall sample error for biplots constructed from 5000 observations from a 5-variate Gamma-Gumbel distribution.	58

5.7	Extreme sample error for biplots constructed from 5000 observations from a 5-variate Gamma-Gumbel distribution.	58
5.8	Overall and extreme sample error simulation results for biplots from a Gamma-Gumbel distribution. Each table consists of the average sample error, standard error and p-values for 4, 5 and 7 variables using 500 and 5000 observations for 20 and 80 tail samples.	60
5.9	Predictivity and Adequacy measures for biplots constructed from 5000 observations from a 4-variate Heterogeneous-Gumbel distribution.	63
5.10	Overall sample error for biplots constructed from 5000 observations from a 4-variate Heterogeneous-Gumbel distribution.	64
5.11	Extreme sample error for biplots constructed from 5000 observations from a 4-variate Heterogeneous-Gumbel distribution.	64
5.12	Overall sample error simulation results for biplots from a 4-variate Heterogeneous-Gumbel distribution. The table constitutes the average sample error, standard error and p-values using 5000 observations and 80 tail samples.	65
5.13	Extreme sample error simulation results for biplots from a 4-variate Heterogeneous-Gumbel distribution. The table constitutes the average sample error, standard error and p-values using 5000 observations and 80 tail samples.	65
6.1	Predictivity and Adequacy measures for biplots constructed using Rand/Foreign currency exchange rate monthly returns	70
6.2	Average overall sample prediction error for Rand/Foreign currency exchange rate monthly returns	70
6.3	Average extreme sample prediction error for Rand/Foreign currency exchange rate monthly returns	71
6.4	Prediction of Rand/Foreign currency exchange rate monthly return on January 2016	73

B.1	Rand foreign currency exchange rate monthly returns for period July 2013 to June 2018.	81
B.2	Standardised Rand foreign currency exchange rate monthly returns for period July 2013 to June 2018.	83

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND AND MOTIVATION

In 1884, Abbott (1884) published a novel titled *Flatland: A romance of many dimensions*. In his novel Abbott (1884) attempted to challenge the notion of conceptualising higher dimensional space. To do this, he chronicles a tale of a square living in a 2-dimensional world known as Flatland. The novel unfolds when the square encounters a sphere living in 3-dimensional space. The sphere struggles to convince the square that he is not a circle, but an object in a higher dimension. The story of Flatland is an excellent thought experiment that demonstrates the difficulty of comprehending higher dimensions. Since humans can only visualise in 3 dimensions, our abilities are constrained when examining higher dimensional phenomena.

This limitation proves to be an obstacle in statistical analysis too, since in the words of Everitt (1994), “*There are many patterns and relationships that are easier to discern in a graphical display than by any other data analysis method*”. To overcome this hindrance, with regards to the visualisation of multivariate data for dimensions higher than three, Gabriel (1971) introduced the biplot. Biplots are a graphical technique constructed using dimension reduction techniques to visualise multivariate data in 2 or 3 dimensions. Note, however, that the “bi” in biplot does not refer to the dimensionality of the display, but due to a biplot displaying both observations and variables, simultaneously.

1.1. BACKGROUND AND MOTIVATION

The biplot was later extended by Gower and Hand (1996) to be used as a multivariate analogue of a scatter plot. There are many methods to construct biplots, however for the purpose of this study, only Principal Component Analysis (PCA) biplots are considered. PCA is a methodology used to reduce the dimensionality of a dataset through identifying principal components, that preserve the maximum variation of the data, in lower dimensions. However, any form of dimension reduction will inevitably lead to a loss of information. This is true for PCA biplots too, however, the loss of information is compensated by the convenience of visualisation. The reason for only considering PCA in this study is twofold. Firstly, PCA is one of the simplest and widely applied dimension reduction techniques. Secondly, PCA and its related methods, such as factor analysis, are widely applied and researched in the field of quantitative finance. In fact, one of the earliest published application of PCA was by Stone (1947), who applied PCA to economic time series data (Jolliffe, 2002). Since then PCA has been used in finance both directly and indirectly, from modelling stock portfolios to analysing and constructing bond curves. Therefore, there is adequate justification that PCA biplots can be applied to visualise multivariate financial data.

Although PCA may be suitable in general for financial data, de Carvalho (2016) argues that PCA may be inappropriate if one's purpose is to analyse the extremes of multivariate data. This poses a problem for financial data since the extremes are essential to characterising risk, whereas the majority of the observations surrounding the mean is of less importance. Since PCA is an inadequate dimension reduction technique for multivariate extremes, it implies directly that PCA biplots may not be well suited for multivariate extremes. To overcome this, Chautru *et al.* (2015) proposes the use of cluster analysis combined with Principal Nested Spheres for dimension reduction of multivariate extremes, but this does not allow for visualisation of the multivariate dataset. The convenience and simplicity of PCA suggests that instead of abandoning PCA altogether, it may be useful to first experiment with adjusting PCA to be more appropriate for extremes. Such an adjustment can be pursued by considering alternative methods to construct covariance and/or correlation matrices. The covariance matrix of a dataset is a critical part of PCA since the PCA methodology is based on preserving the maximum variation of a dataset. Similarly, if each variable in a multivariate dataset is

1.1. BACKGROUND AND MOTIVATION

scaled by its standard deviation and shifted to have a mean of zero, then the covariance and correlation matrix are the same. Hence, PCA can be executed on a correlation matrix of a scaled dataset. Jolliffe (2002) endorses PCA on a correlation matrix instead of a covariance matrix for two reasons. The first is that if variables are measured using different units, then PCA using a covariance matrix will be biased towards preserving the variable with the largest variation on its particular measurement scale. Secondly, it is more difficult to compare the results of PCA for different analyses when using the covariance matrix. However, the use of a covariance matrix does have its own advantages. The first being that if inference is the goal of statistical analysis then PCA on a covariance matrix is superior. Additionally, it is not possible to unscale PCA estimated data to represent the original data set when using the correlation matrix. Nonetheless, since the objective of PCA biplots is visualisation, PCA on the correlation matrix is used throughout this study to construct PCA biplots.

Owing to the fact that correlations will be used to perform PCA, the question is then how can the correlation matrix be adjusted to accommodate for multivariate extremes? It is firstly important to acknowledge that correlation is not the only way to measure dependence. Furthermore, correlation as a measure of dependence has the disadvantage of only measuring linear dependence. Moreover, Klüppelberg and Stelzer (2014) argues that in the context of risk one does not really care about correlation since the correlation depends on the whole distribution. Whereas, it is of more value in the risk management setting to find the dependence of extreme outcomes. Therefore a more rigorous approach to characterising dependence and more specifically extreme dependence is required. A possible solution is to consider copulas. According to Nelsen (2007), copulas are functions that join multivariate distribution functions from univariate marginal distribution functions. Given that copulas link variables, it therefore fully describes the dependence of the underlying variables in a multivariate distribution. For this reason, this study will investigate how copulas can be used to construct correlation matrices. More specifically, the ability to construct a correlation matrix using copulas for multivariate extreme observations. Given that multivariate extremes are the main concern, an appropriate dimension reduction technique would aim to preserve extremal dependencies instead of maximum variation. If a suitable adjust-

1.1. BACKGROUND AND MOTIVATION

ment for PCA is found to accommodate extremes, it will as a consequence, improve PCA biplots for multivariate extremes. An approach with this in mind was developed by Haug *et al.* (2015) who used an elliptical copula that is calibrated using the tail dependence function to construct a correlation matrix for multivariate extreme observations. It can then be argued that performing PCA on a correlation matrix constructed from extremes should be more inclined to preserve extreme observations. Hence, an investigation into the application of such an extremal correlation matrix to construct PCA biplots is the primary objective of this study.

The study is performed by undertaking the following objectives:

1. A detailed discussion on the background and construction of PCA biplots.
2. Investigate the various PCA biplot quality measures.
3. Introduce the concept of dependence and various techniques to measure dependence.
4. Study in detail the development, use and properties of several copula families.
5. Demonstrate how elliptical copulas can be used to determine a correlation matrix for multivariate extremes.
6. Propose the use the derived correlation matrix for multivariate extremes, to improve PCA biplots for extreme observations.
7. Implement a simulation study to evaluate whether the proposed methodology improves PCA biplots for extreme observations.
8. Illustrate the suitability of the improved PCA biplot on real-world data that is subject to multivariate extremes.

To achieve the above-mentioned objectives this study is presented in the following sequence:

In **Chapter 2** the necessary background for PCA biplots is discussed in detail. The chapter starts by firstly introducing the PCA methodology. This is followed by a demonstration on how PCA is used to construct PCA biplots. By way of an example, the use

1.1. BACKGROUND AND MOTIVATION

and interpretation of a PCA biplots are then illustrated. The chapter ends by providing some PCA biplot quality measures that are used to assess the overall fit of the biplot.

Chapter 3 is dedicated to exploring the idea behind measuring dependence. It starts by introducing various dependence measures and discusses the purpose of each measure. The main component of this chapter is a review of copulas to measure dependence. This is done by discussing the theory underlying copulas as well as in-depth look into two copula families namely, the Elliptical and Archimedean copula families. It ends by reviewing a noteworthy application of copulas used to determine correlation matrices for multivariate extreme observations.

In **Chapter 4**, the literature discussed in Chapter 2 and 3 is used to develop a methodology to improve PCA biplots for extreme observations. This constitutes a discussion regarding the approach taken to improve PCA biplots as well as how such an improvement is evaluated.

In **Chapter 5**, an empirical study by way of simulation is pursued, in order to determine if the improved PCA biplot for extremes performs better than the traditional PCA biplots when assessing the fit of extreme values. This is done by simulating observations from several multivariate distributions and examining the biplot fit at extremes for the improved and the traditional biplot methodology.

Chapter 6 is a short chapter whereby the improved biplot methodology is tested against the traditional biplot methodology on a real-world financial dataset.

Finally, in **Chapter 7**, the work carried out and contributions made are summarised and areas of further research are suggested.

CHAPTER 2

A REVIEW OF PCA BILOTS

As stated in the previous chapter, the use of PCA biplots (subsequently referred to as biplots) will be the focus of this chapter. The purpose of this chapter is to provide some needed background on the derivation and interpretation of biplots. Note that, some mathematical background required for this chapter is provided in Appendix A. Firstly, some background on PCA is given, then the use of PCA to construct biplots is discussed. This is followed by an explanation of how biplots can be interpreted by way of an example. The final section presents some useful techniques to measure the quality of a biplot display.

2.1 PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA is one of the most popular dimension reduction techniques. The popularity of PCA is due to its simplicity and due to it being widely researched and applied in many fields of study. PCA originates from publications by Pearson (1901) and Hotelling (1933), who independently derived PCA using differing approaches. The differences in their approaches are owed to them having different motivations for their use of PCA. Pearson (1901) was concerned with finding some lines and planes that best fit observations in a p -dimensional space. Hotelling (1933), on the other hand, wanted to determine observations of p variables by finding some smaller set of independent variables, similar to the idea of factor analysis (Jolliffe, 2002). For the purpose of this study, when

2.1. PRINCIPAL COMPONENT ANALYSIS (PCA)

discussing PCA biplots, PCA will be viewed from the perspective of Pearson (1901).

The first step in PCA is to derive the principal components (PCs) of the underlying data matrix. PCs are derived by minimising the sum of the squared orthogonal distances (residuals) between the original p -variable space and the reduced r -dimensional subspace. Further, using the Huygens Principle, Gower and Hand (1996) proved that the optimal r -dimensional subspace is one that passes through the centroid of points $\bar{\mathbf{X}}$ in the p -variable space. This means that in order to optimally reduce a p -dimensional space to an r -dimensional space, the underlying observations \mathbf{X} should be centred. Therefore, throughout this study, the assumption is made that an underlying data matrix \mathbf{X} is preprocessed to be centred, i.e. $E[\mathbf{X}] = [\mu_1, \dots, \mu_p]' = [0, \dots, 0]$.

Suppose $\mathbf{X} : n \times p$ is a centred data matrix with p variables and n observations where the observation is denoted by the vector $\mathbf{x}_i, i = 1, 2, \dots, n$. Then $\mathbf{X}'\mathbf{X}$ is proportional to the sample covariance matrix of \mathbf{X} , which can be presented by applying Singular Value Decomposition (SVD)¹ as:

$$\mathbf{X}'\mathbf{X} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}' \quad (2.1)$$

where, $\mathbf{\Lambda} : p \times p$ is a diagonal matrix containing the ordered (from largest to smallest) eigenvalues of $\mathbf{X}'\mathbf{X}$, denoted $\lambda_i, i = 1, 2, \dots, p$, and $\mathbf{V} : p \times p$ is a matrix containing the orthonormal eigenvectors of $\mathbf{X}'\mathbf{X}$ as its column vectors, ordered accordingly. The p column vectors of \mathbf{V} denoted $\mathbf{v}_i, i = 1, 2, \dots, p$ are termed the sample principal components (Sample PCs). Further, the matrix of principal component scores (PC Scores), $\mathbf{Z} : n \times p$, is determined as:

$$\mathbf{Z} = \mathbf{X}\mathbf{V} \quad (2.2)$$

which are coordinates of the sample PCs in the p -dimensional space.

To reduce the dimensionality of the data matrix to r -dimensional space the first r column vectors are extracted and denoted as $\mathbf{V}_r : p \times r$. Thus, \mathbf{V}_r is a matrix containing the first r eigenvectors $\mathbf{v}_i, i = 1, 2, \dots, r$ corresponding the r largest eigenvalues. It then follows that the principal component approximation of \mathbf{X} is given by:

$$\hat{\mathbf{Z}} : n \times p = \mathbf{X}\mathbf{V}_r\mathbf{V}_r' \quad (2.3)$$

¹See Appendix A.2 for a detailed explanation of singular value decomposition (SVD)

2.2. PCA BIPLLOT CONSTRUCTION

this approximation yields the smallest sum of squared residuals between the original observations in p -dimensional spaces and its projection in r -dimensional space, i.e. $\|\mathbf{X} - \hat{\mathbf{Z}}\|^2$ is minimised.

Alternatively, PCA can be performed by applying SVD directly on the data matrix \mathbf{X} . The SVD of \mathbf{X} is given by:

$$\mathbf{X} = \mathbf{U}\mathbf{\Omega}\mathbf{V}' \quad (2.4)$$

Since \mathbf{V} is an orthogonal matrix, multiplying by \mathbf{V} on the right of (2.4) yields,

$$\mathbf{X}\mathbf{V} = \mathbf{U}\mathbf{\Omega} = \mathbf{Z} \quad (2.5)$$

which is the PC scores as in (2.2). To find the r -dimensional subspace the largest r singular values of $\mathbf{\Omega}$ is extracted and denoted as $\mathbf{\Omega}_r$. Correspondingly, let \mathbf{U}_r and \mathbf{V}_r be denoted as the first r columns of \mathbf{U} and \mathbf{V} , respectively. Then it follows that the best r -dimensional approximation of the data matrix \mathbf{X} is given as:

$$\mathbf{X} \approx \mathbf{X}\mathbf{V}_r\mathbf{V}_r' = \mathbf{U}_r\mathbf{\Omega}_r\mathbf{V}_r' \quad (2.6)$$

The use of the derived sample PCs and PC scores to construct PCA biplots is discussed in the next section.

2.2 PCA BIPLLOT CONSTRUCTION

As stated in Chapter 1, Gabriel (1971) introduced PCA biplots as a way to jointly represent the observations and variables of a dataset. Then Gower and Hand (1996) adjusted the PCA biplot to represent a multivariate version of the traditional scatter plot which is used throughout this study. The construction of the PCA biplot is the focus of this section.

Gabriel (1971) states that the decomposition in (2.6) can also be represented as:

$$\mathbf{X} \approx \mathbf{U}_r\mathbf{\Omega}_r^\alpha\mathbf{\Omega}_r^{1-\alpha}\mathbf{V}_r' = (\mathbf{U}_r\mathbf{\Omega}_r^\alpha)(\mathbf{V}_r'\mathbf{\Omega}_r^{1-\alpha})' = \mathbf{G}_r\mathbf{H}_r \quad (2.7)$$

with $\mathbf{G}_r = \mathbf{U}_r\mathbf{\Omega}_r^\alpha$, $\mathbf{H}_r' = \mathbf{V}_r'\mathbf{\Omega}_r^{1-\alpha}$, and $0 \leq \alpha \leq 1$. The purpose and use of a biplot display is determined by the value of α . This can be explained by separately considering the

2.2. PCA BIPLLOT CONSTRUCTION

case for $\alpha = 0$ and $\alpha = 1$. If $\alpha = 0$, then $\mathbf{G}_r = \mathbf{U}_r$ and $\mathbf{H}'_r = \mathbf{\Omega}_r \mathbf{V}_r$ since \mathbf{U}_r is orthonormal the following can be derived:

$$\mathbf{X}'\mathbf{X} \approx (\mathbf{G}_r \mathbf{H}'_r)'(\mathbf{G}_r \mathbf{H}'_r) = \mathbf{H}_r \mathbf{G}'_r \mathbf{G}_r \mathbf{H}_r = \mathbf{H}_r \mathbf{H}_r. \quad (2.8)$$

Therefore, entries in \mathbf{H}_r uniquely determine the covariance matrix of \mathbf{X} . Additionally, it can be proven that when $\alpha = 0$ the Mahalanobis distances are approximated between observations. Alternatively, if $\alpha = 1$ then $\mathbf{G}_r = \mathbf{U}_r \mathbf{\Omega}_r$ and $\mathbf{H}'_r = \mathbf{V}'_r$ and it can be proven in this case that the euclidean distances are approximated between observations. The differences between Mahalanobis² and euclidean distances is beyond the scope of this study. However, the choice of $\alpha = 1$ will result in observations being represented better than variables on the biplot and if $\alpha = 0$, it results in the variables, and as a consequence correlations, being represented better than the observations on the biplot. The PCA biplot of Gower and Hand (1996) assigns $\alpha = 1$, which is the assumption for the rest of this chapter. The biplot in the case where $\alpha = 0$ is referred to as correlation biplot since variables are better represented.

In the previous section, it was shown that the best r -dimensional subspace to represent observations from a p -dimension space is determined by the first r eigenvectors (Sample PCs) of $\mathbf{X}'\mathbf{X}$ denoted by \mathbf{V}_r . The columns of \mathbf{V}_r provide a set of orthogonal coordinate axes in the r -dimensional space termed the principal axes. The principal axes are used only for representing the biplot observations and is also referred to as the scaffolding axes. The biplot observations are determined as projections from the principal axes and are given by,

$$\mathbf{Z}_r = \mathbf{X}\mathbf{V}_r \quad (2.9)$$

where, the rows of \mathbf{Z}_r represents the PC scores for the first r sample PCs and is denoted as $\mathbf{z}_i : i = 1, 2, \dots, n$.

The next step in the biplot construction is deciding whether the biplot will be used for interpolation or prediction. In the case of interpolation a new p -variable observation $\mathbf{x}^* : p \times 1$ has to be projected to an observation in the r -dimensional space as $\mathbf{z}^* : p \times 1$.

²This is a generalised method to measure distance, introduced by Mahalanobis (1936). The Mahalanobis distance measures the number of standard deviations a point is from the mean of its distribution.

2.3. PCA BIPLLOT INTERPRETATION

This r -dimension projection can be obtained using (2.9) as

$$\mathbf{z}^{*'} = \mathbf{x}^{*'} \mathbf{V}_r \quad (2.10)$$

Alternatively, in the case of prediction the original p -variable observation must be approximated as $\hat{\mathbf{x}}^* : p \times 1$ from the coordinates in the r -dimensional space \mathbf{z}^* . This can be found using (2.6) as

$$\hat{\mathbf{x}}^* = \mathbf{z}^{*'} \mathbf{V}_r \mathbf{V}_r' \quad (2.11)$$

The choice between interpolation and prediction is not trivial since the biplot axes markers are different in both cases. Therefore, if the purpose of a biplot display is for both prediction and interpolation the two separate biplots must be constructed. However, for the purposes of this study only predictive axes will be used.

The final step in biplot construction is plotting the axes that correspond to the p -variables of the data. As stated, axes for prediction and interpolation will differ in terms of the position of the axes markers. The different axes markers are determined by some value of μ , with $-\infty < \mu < \infty$. Suppose $\mathbf{e}_k : r \times 1$ is a unit vector with the k^{th} element equal to one and all other elements equal to zero. Then each observation \mathbf{x}_i with coordinates $(x_{i,1}, x_{i,2}, \dots, x_{i,p})$ can be written as,

$$\mathbf{x}_i = \sum_{k=1}^p x_{i,k} \mathbf{e}_k \quad (2.12)$$

this will interpolate to the point,

$$\mathbf{z}_i' = \mathbf{x}_i' \mathbf{V}_r = \sum_{k=1}^p x_{i,k} \mathbf{e}_k' \mathbf{V}_r \quad (2.13)$$

Therefore the k^{th} interpolation biplot axis markers is determined by $\mu \mathbf{e}_k \mathbf{V}_r$. It can further be shown that the corresponding k^{th} prediction biplot axis markers is given by,

$$\frac{\mu \mathbf{e}_k \mathbf{V}_r}{\mathbf{e}_k \mathbf{V}_r \mathbf{V}_r' \mathbf{e}_k'} \quad (2.14)$$

as μ varies.

2.3 PCA BIPLLOT INTERPRETATION

This section explores the interpretation of a biplot using an example of data from the risk management field. The example used in this section was initially presented by

2.3. PCA BIPLLOT INTERPRETATION

Van Blerk (2000) and further discussed in the textbook by Gower *et al.* (2011). The data used is daily 95% Value-at-Risk (VaR) observations for 7 financial trading desks over a 20 day period. The data used is provided in Table 2.1 and consists of 20 observations for 7 variables. Biplots are constructed throughout this study using the R package `UBbipl` developed by le Roux and Lubbe (2013). The R code used to obtain the biplots and results presented below is given in Appendix C.1. The biplot for the 95% VaR dataset is illustrated in Figure 2.1. The axes in the biplot represent the variables which are the 95% VaR values for each of the 7 financial trading desks and the VaR observations for the portfolio are presented by the green points each labelled corresponding to the day of the measurement.

Table 2.1

Van Blerk (2000) 95% VaR of financial trading desks

Day	CM	IRD	MM	ALCO	SE	EDSA	EDM
1	-1.7647	-0.2481	-0.2810	-0.2961	-0.1406	-0.2262	-0.9409
2	-0.8181	-1.3258	-0.2810	-0.2961	-0.1419	0.0123	-3.3836
3	-1.7152	-1.1400	-0.5961	-0.2961	-0.1410	-0.1825	-2.8719
4	-1.7714	-1.6412	-0.5961	-0.2961	-0.1454	-0.8900	-1.9459
5	-1.6613	-1.3016	-0.4124	-0.4755	-0.1319	-0.2153	-1.2899
6	0.0219	-1.3635	-0.6078	-0.2789	-0.2155	-0.2987	-1.3775
7	-0.8892	-1.1370	-0.4568	-0.4531	-0.1523	-0.2549	-1.1285
8	-0.9138	-1.1991	-0.4568	-0.4041	-0.1466	-0.0834	-1.1372
9	-1.1491	-1.1821	-0.4568	-0.4041	-0.1489	-0.3568	-1.1747
10	-1.2728	-0.7334	-0.4568	-0.4041	-0.1565	-0.5556	-0.8941
11	-0.8168	-0.8515	-0.4568	-0.4041	-0.1667	-0.3794	-0.8884
12	-1.2067	-1.5127	-0.4568	-0.4568	-0.1613	-0.0376	-0.8037
13	-0.8625	-1.8187	-0.4592	-0.4568	-0.1577	-0.1392	-0.9391
14	-2.5521	-1.4004	-0.4592	-0.4568	-0.1651	-0.1398	-0.9136
15	-1.4310	-1.3198	-0.4592	-0.4568	-0.1684	-0.1373	-1.0968
16	-2.8378	-1.3177	-0.4592	-0.4568	-0.1584	-0.3692	-0.1620
17	-1.0766	-1.2734	-0.7296	-0.4568	-0.2560	-0.0889	-1.1253
18	-1.0256	-1.3378	-0.7296	-0.4357	-0.2774	-0.2957	-1.0238

Continued on next page

2.3. PCA BIPLLOT INTERPRETATION

Table 2.1 – continued from previous page

Day	CM	IRD	MM	ALCO	SE	EDSA	EDM
19	-1.0462	-1.3070	-0.0253	-0.0311	-0.1352	-0.3648	-0.6462
20	-0.6270	-2.0298	-0.0246	-0.0311	-0.1318	-0.1381	-0.6129

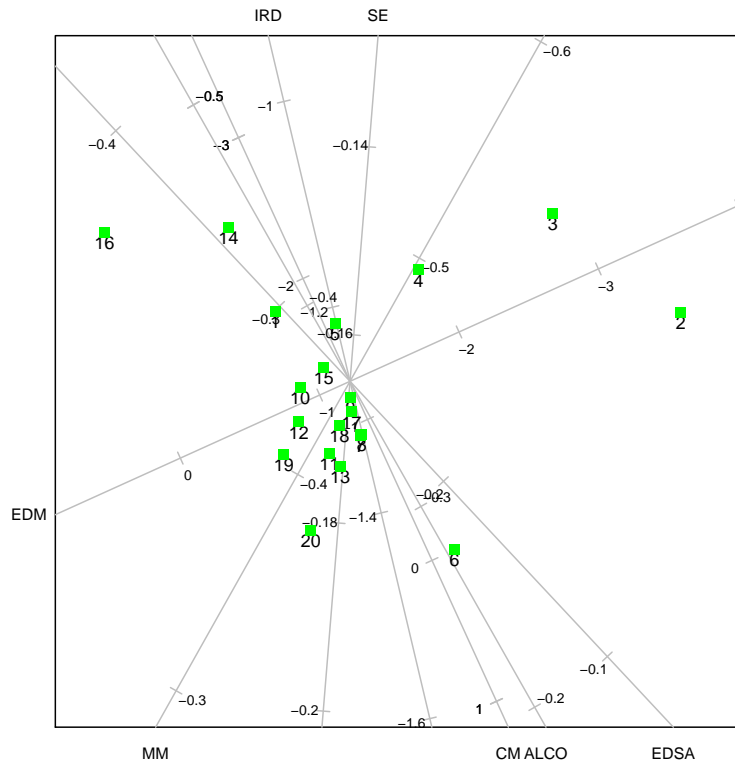


Figure 2.1

Biplot for 95% VaR data.

The angles between biplot axes give an approximation of the correlation between variables, where a small angle between axes indicates that variables are highly correlated and orthogonal axes indicate that observations have low correlation. However, the biplot constructed in Figure 2.1 takes $\alpha = 1$ as in (2.7), meaning that observations are better represented than variables. Conversely, in Figure 2.2 the correlation biplot is constructed with $\alpha = 0$, which better approximates the correlations between variables by the angles between axes since variables are better represented than observations.

2.3. PCA BIPLLOT INTERPRETATION

When comparing Figures 2.1 and 2.2 with Table 2.2 providing the correlation matrix of the trading desk VaR values, it can be seen that neither biplot perfectly represents the correlation of the trading desks. However, Figure 2.2 does slightly better in representing correlations, for example the angle between **ALCO** and **EDSA** is slightly larger in Figure 2.2 to account for its low correlation.

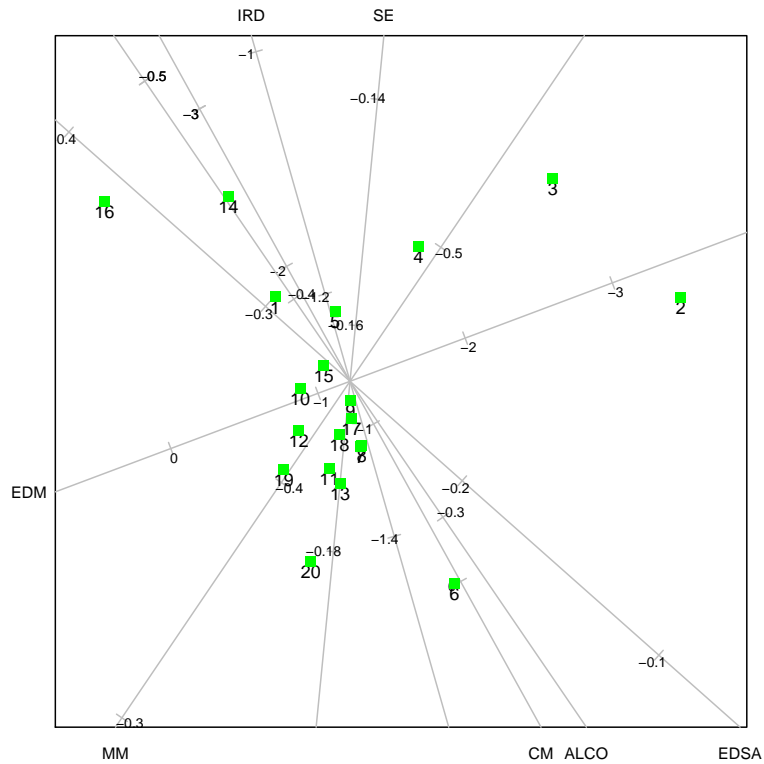


Figure 2.2

Correlation biplot for 95% VaR data.

2.3. PCA BIPLLOT INTERPRETATION

Table 2.2

Correlation matrix for 95% VaR of 7 financial trading desks

	CM	IRD	MM	ALCO	SE	EDSA	EDM
CM	1.00	-0.16	0.08	0.31	-0.23	0.18	-0.14
IRD	-0.16	1.00	-0.08	-0.19	0.03	-0.11	-0.01
MM	0.08	-0.08	1.00	0.69	0.68	0.15	0.17
ALCO	0.31	-0.19	0.69	1.00	0.32	-0.13	-0.09
SE	-0.23	0.03	0.68	0.32	1.00	-0.07	-0.12
EDSA	0.18	-0.11	0.15	-0.13	-0.07	1.00	-0.10
EDM	-0.14	-0.01	0.17	-0.09	-0.12	-0.10	1.00

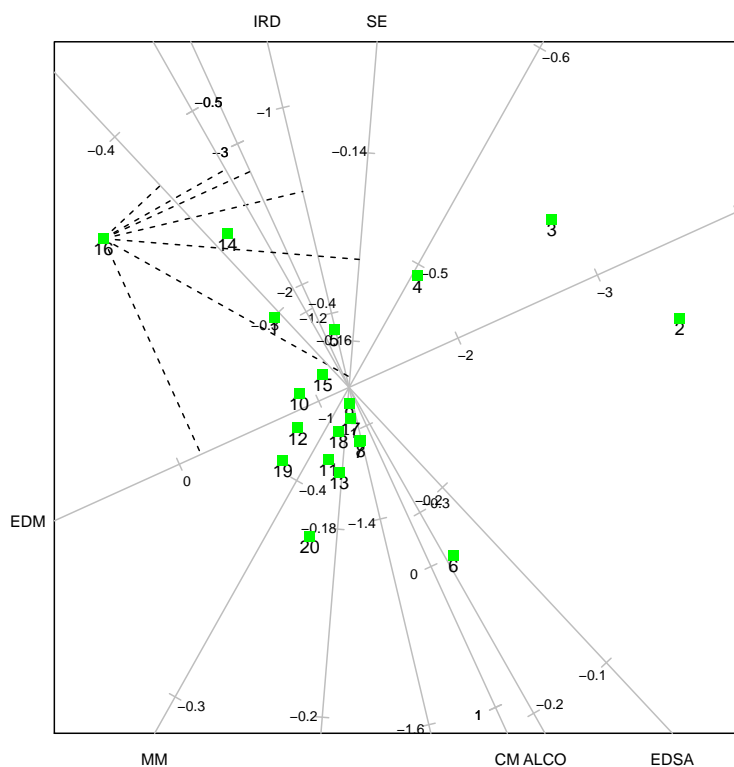


Figure 2.3

Biplot for 95% VaR data with predictions of day 16 VaR.

2.4. PCA BIPLLOT QUALITY MEASURES

Table 2.3

Actual and predicted value of the 95% VaR for day 16

	CM	IRD	MM	ALCO	SE	EDSA	EDM
Actual	-2.838	-1.318	-0.459	-0.457	-0.158	-0.369	-0.162
Predicted	-2.803	-1.082	-0.447	-0.471	-0.151	-0.372	-0.156

To read off observations from the biplot, orthogonal lines are drawn from the observation to each of the axes as illustrated in Figure 2.3, for the VaR sample on day 16. Since the biplot is a 2-dimensional approximation of observations in 7-dimensional space, there is a disparity between the actual observation values and those approximated on the biplot. The difference in the biplot approximation of the VaR on day 16 and its actual value is presented in Table 2.3. Further, some variables are better approximated than others for example the discrepancy in **IRD** is large compared to the other instruments. Thus, given that a biplot is a visual approximation, measures of how well the biplot displays the underlying data are required and is discussed in the next section.

2.4 PCA BIPLLOT QUALITY MEASURES

In the preceding sections, the construction and interpretation of a biplot were discussed. However, the biplot interpretation is meaningless if the biplot does not display the data reasonably. Therefore, in this section, some biplot fit quality measures are explored. The quality measures considered in this section are presented in more detail in the master's thesis by Brand (2013).

Suppose throughout this section that \mathbf{X} is a centred data matrix whose i^{th} observation is denoted as \mathbf{x}_i . Then from (2.6), \mathbf{X} is approximated by,

$$\hat{\mathbf{X}} = \mathbf{X} \mathbf{V}_r \mathbf{V}_r' \quad (2.15)$$

The first measure of fit will be a measure of the overall quality of the biplot representation denoted as \mathcal{V} . The overall quality is determined as the ratio of the fitted sum of

2.4. PCA BIPLLOT QUALITY MEASURES

squares and the total sum of squares as,

$$\mathcal{V} = \frac{\text{trace}\{\hat{\mathbf{X}}'\hat{\mathbf{X}}\}}{\text{trace}\{\mathbf{X}'\mathbf{X}\}}. \quad (2.16)$$

Now, (2.16) can be further simplified using (2.15) and (A.3) as:

$$\begin{aligned} \mathcal{V} &= \frac{\text{trace}\{\hat{\mathbf{X}}'\hat{\mathbf{X}}\}}{\text{trace}\{\mathbf{X}'\mathbf{X}\}} \\ &= \frac{\text{trace}\{\mathbf{V}_r\boldsymbol{\Omega}_r^2\mathbf{V}_r'\}}{\text{trace}\{\mathbf{V}\boldsymbol{\Omega}^2\mathbf{V}'\}} \\ &= \frac{\text{trace}\{\boldsymbol{\Omega}_r^2\mathbf{V}_r'\mathbf{V}_r\}}{\text{trace}\{\boldsymbol{\Omega}^2\mathbf{V}'\mathbf{V}\}} \\ &= \frac{\text{trace}\{\boldsymbol{\Omega}_r^2\}}{\text{trace}\{\boldsymbol{\Omega}^2\}} \\ &= \frac{\sum_{i=1}^r \lambda_i}{\sum_{i=1}^p \lambda_i} \end{aligned} \quad (2.17)$$

where, λ_i is the i^{th} largest eigenvalue of $\mathbf{X}'\mathbf{X}$. Given that $\mathbf{X}'\mathbf{X}$ is positive semi-definite it implies that $\lambda_i \geq 0, \forall i$ thus $0 \leq \mathcal{V} \leq 1$. Overall quality is at its maximum if $\hat{\mathbf{X}} = \mathbf{X}$. It is important to note that, even if the overall biplot quality is low, it can still be possible that some individual observations and variables are reasonably presented.

The next quality measure assesses how well the biplot axes represent the axes of the variables in the p -dimensional space which is termed the adequacy of the axes. The adequacy for the axis representing the k^{th} variable is denoted as γ^k and is given by Gardner-Lubbe *et al.* (2008) as

$$\gamma^k = \sum_{j=1}^r v_{kj}^2 \quad (2.18)$$

where, v_{kj}^2 is the k^{th} row and j^{th} column entry of the matrix $\mathbf{V}_r\mathbf{V}_r'$.

Another quality measure proposed by Gardner-Lubbe *et al.* (2008) measures the predictive ability of a biplots axes. The overall predictability of the biplots axes is expressed as:

$$\Pi = \frac{\text{diag}\{\hat{\mathbf{X}}'\hat{\mathbf{X}}\}}{\text{diag}\{\mathbf{X}'\mathbf{X}\}}. \quad (2.19)$$

Further, it can be shown that the predictivity of the k^{th} biplot axis is given by,

$$\pi_k = \frac{\sum_{i=1}^n \hat{x}_{ik}^2}{\sum_{i=1}^n x_{ik}^2}. \quad (2.20)$$

It is further shown by Gardner-Lubbe *et al.* (2008) that the sum of the axis predictivity weighted by the variance of the underlying variable is equal to the overall quality of the

2.5. SUMMARY

biplot. Thus, high overall quality does not imply that axis predictivity is high for each of the p variables.

The final biplot quality measure is presented by Brand (2013) and is termed the sample predictivity. The sample predictivity measures the accuracy at which samples are approximated by the biplot. This implies that, if an observation has a low sample predictivity, then relative sample positions on the biplot is questionable. The sample predictivity for the i^{th} sample denoted as ψ_i is expressed as,

$$\psi_i = \frac{\hat{\mathbf{x}}_i' \hat{\mathbf{x}}_i}{\mathbf{x}_i' \mathbf{x}_i}. \quad (2.21)$$

Similarly, the sample predictivity can also be evaluated for the i^{th} sample as

$$\psi_i^* = (\hat{\mathbf{x}}_i - \mathbf{x}_i)^2. \quad (2.22)$$

The sample prediction obtained using (2.22) will provide the sample prediction error for each variable and can be summed to find the overall sample predictivity.

All measures presented in this section are essential to the biplot display, as it presents a way to assess whether the biplot is a reasonable visualisation of the underlying data. Further, these measures should always be evaluated when a biplot is used since it provides assurance as to whether the biplot representation is realistic.

2.5 SUMMARY

In this chapter, some necessary background on PCA and its application to construct biplots were provided. Further, an example of a biplot was presented accompanied by guidance on how to interpret biplots. Finally, some noteworthy methods to assess the quality of a biplot display was discussed. The next chapter will provide some background on dependence measures and copulas which are later used to improve the quality of a biplot.

CHAPTER 3

A REVIEW OF COPULAS AND DEPENDENCE

An essential part of multivariate data analysis is to identify relationships between variables, which is characterised by the underlying dependence between variables. The importance of dependence is extended by Jogdeo (1982) who noted, “*Dependence relations between random variables is one of the most widely studied subject in probability and statistics. The nature of dependence can take a variety of forms and unless some specific assumptions are made about dependence, no meaningful statistical model can be contemplated*”. In the previous chapter, linear correlation, as a measure of dependence, was used to perform PCA. This means that the concept of dependence is essential to PCA biplot construction. This chapter is therefore an examination of how to measure and interpret dependence as well as what underlying assumptions are required to characterise dependence. This chapter starts with some background on traditional measures of dependence. This is followed by a more modern look at dependence through the use of copula functions. Finally, this chapter is concluded by examining how to evaluate dependence for multivariate extremes.

3.1. DEPENDENCE MEASURES

3.1 DEPENDENCE MEASURES

In order to understand what it entails for variables to be dependent it is necessary to first define what it means for variables to be independent. Independence can be defined by considering a set of d random variables X_1, X_2, \dots, X_d . Then this set of variables are independent if and only if,

$$P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_d \leq x_d) = P(X_1 \leq x_1)P(X_2 \leq x_2) \cdots P(X_d \leq x_d) \quad (3.1)$$

If (3.1) does not hold it can be assumed that there exists some level of dependence between the underlying random variables. There are, however, many ways to measure dependence between variables and some of these dependence measures are discussed in this section.

The most well known and widely applied dependence measure is Pearson's correlation coefficient (ρ), which is a measure of linear dependence between two random variables. Suppose that X, Y is a pair of random variables with some linear relationship. Then for these two variables, $\rho(X, Y)$ can be calculated as,

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}, \quad (3.2)$$

with $-1 \leq \rho(X, Y) \leq 1$. If X and Y are independent then $\rho(X, Y) = 0$. However, the converse does not hold, in other words, $\rho(X, Y) = 0$ does not imply independence. Further, X and Y have perfect linear dependence if $\rho(X, Y) = \pm 1$. The popularity of Pearson's correlation coefficient as a measure of dependence is mainly due to it being easy to estimate from observed data. Additionally, Pearson's correlation serves as a natural dependence measure if data is distributed multivariate normally and as such, only in the case of multivariate normally distributed data does $\rho = 0$ imply independence. However, there are many disadvantages of Pearson's correlations coefficient. The biggest disadvantage is that it only measures linear dependence and that it is not invariant under monotone transformations.¹

The next measures of dependence that overcome some of the disadvantages of Pearson's correlation are measures of rank dependence or also referred to as ordinal de-

¹Further, correlation disadvantages can be found in Embrechts *et al.* (2002).

3.1. DEPENDENCE MEASURES

pendence. These dependence measures do not consider the magnitude of the observations, but only the order of the observations. The first ordinal dependence measure discussed is Spearman's correlation coefficient (ρ_S), which is sometimes also referred to as Pearson's correlation for ranked variables (Meissner, 2013). To define ρ_S , let X and Y be continuous random variables with distribution functions F_1 and F_2 , respectively. Then Spearman's correlation coefficient is given by,

$$\rho_S(X, Y) = \rho(F_1(X), F_2(Y)) \quad (3.3)$$

where ρ is Pearson's correlation coefficient as in (3.2). Since $F_1(X)$ and $F_2(Y)$ are standard uniform random variables the magnitudes of X and Y are irrelevant and thus only orders of X and Y are considered. Now, Klüppelberg and Stelzer (2014) states that ρ_S can be calculated empirically as,

$$\hat{\rho}_S(X, Y) = \frac{1}{2}n(n^2 - 1) \sum_{i=1}^n \left[\text{rank}(x_i) - \frac{n+1}{2} \right] \left[\text{rank}(y_i) - \frac{n+1}{2} \right] \quad (3.4)$$

where, n is the number of observations and $(\text{rank}(x_i), \text{rank}(y_i))$ are the ranks of the observations of (X, Y) . Additionally, $-1 \leq \rho_S \leq 1$ with the same interpretation as for Pearson's correlation.

The second ordinal dependence measure discussed was derived by Kendall (1938) and is known as Kendall's tau (τ). Suppose that (X_1, Y_1) and (X_2, Y_2) are random vectors with bivariate distribution function $F(X, Y)$. Then Kendall's tau is given by,

$$\tau(X, Y) = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]. \quad (3.5)$$

However, it is easier to grasp Kendall's tau by considering its empirical version. The empirical formula for Kendall's tau requires the definition of concordant and discordant observation pairs. The pair of observations (X_i, Y_i) and (X_j, Y_j) are said to be:

- i. concordant, if both $X_i > X_j$ and $Y_i > Y_j$, or if both $X_i < X_j$ and $Y_i < Y_j$.
- ii. discordant, if $X_i > X_j$ and $Y_i < Y_j$, or if $X_i < X_j$ and $Y_i > Y_j$.
- iii. neither concordant or discordant, if $X_i = X_j$ and $Y_i = Y_j$.

3.1. DEPENDENCE MEASURES

The number of concordant and discordant pairs is then counted and the empirical Kendall's tau is calculated as,

$$\begin{aligned}\hat{\tau} &= \frac{(\text{Number of concordant pairs}) - (\text{Number of discordant pairs})}{\text{Total number of pairs}} \\ &= \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \text{sign}[(X_i - X_j)(Y_i - Y_j)]\end{aligned}\quad (3.6)$$

with, $\text{sign} = 1$ if observation pairs are concordant or $\text{sign} = -1$ if observation pairs are discordant. When $\hat{\tau} = 1$ it implies that an increase in X always coincides with an increase in Y and vice versa. The main advantage of ordinal dependence measures over linear dependence measures is that ordinal measures are invariant to monotone increasing transformations.

The final dependence measure discussed in this section pertains to a measure of dependence at multivariate extremes. Extreme dependence is quantified through a tail dependence function. Naturally, a tail dependence function measures the dependency of the data in the tail. A tail dependence function distinguishes the dependency in the tail of the data to that of the regular data. Tail dependence is measured for a pair of random variables (X, Y) by determining if there is a non-zero probability that X is large given that Y is large. Hence, random variables (X, Y) are said to be tail independent if

$$\lim_{t \rightarrow \infty} P(Y > t | X > t) \rightarrow 0 \quad (3.7)$$

Conversely, if the limit in (3.7) is non-zero then (X, Y) is said to be tail dependent. The tail dependence function characterises the strength of the dependence in the upper- and/or lower tail of a multivariate distribution. Therefore, it is not necessarily the case that the tail dependence in the upper- and lower tail of the distribution is identical and as such upper- and lower tail dependence is measured individually. The tail dependence function is defined by Klüppelberg *et al.* (2007) as follows,

Definition 3.1 (Tail dependence function). *Suppose $\mathbf{X} = (X_1, \dots, X_d)'$ is a random vector with distribution function F and continuous marginals F_1, \dots, F_d . Define the upper tail dependence function of \mathbf{X} as*

$$\lambda_{upper}^{\mathbf{X}}(u_1, \dots, u_d) = \lim_{t \rightarrow \infty} \frac{1}{t} P[1 - F_1(X_1) \leq tu_1, \dots, 1 - F_d(X_d) \leq tu_d] \quad (3.8)$$

with $(u_1, \dots, u_d) \in [0, 1]^d$.

3.2. COPULA THEORY

Similarly, define the lower tail dependence function of \mathbf{X} as,

$$\lambda_{lower}^X(u_1, \dots, u_d) = \lim_{t \rightarrow \infty} \frac{1}{t} P[F_1(X_1) \leq tu_1, \dots, F_d(X_d) \leq tu_d] \quad (3.9)$$

with $(u_1, \dots, u_d) \in [0, 1]^d$. ■

Now, $\lambda_{upper}^X(1, 1, \dots, 1)$ is termed the upper tail dependence coefficient and $\lambda_{lower}^X(0, 0, \dots, 0)$ is termed the lower tail dependence coefficient. Finally, Klüppelberg *et al.* (2007) defines an empirical tail dependence function as follows:

Definition 3.2. Suppose $\mathbf{X} = (X_1, \dots, X_d)'$ is a random vector containing n samples, with $\mathbf{x}_h = (x_{h,1}, \dots, x_{h,d})'$ for $h = 1, \dots, n$, we define an empirical tail dependence function for $\mathbf{x} > \mathbf{0}$ as,

$$\lambda^{emp}(\mathbf{x}; k) := \frac{1}{k} \sum_{h=1}^n \mathbf{I} \left\{ 1 - F_j(x_{h,j}) \leq \frac{k}{n} x_j \right\} \quad j = 1, \dots, d \quad (3.10)$$

where $1 \leq k \leq n$ and F_j denotes the empirical distribution function of $\mathbf{X}_j, j = 1, \dots, d$. Furthermore, define the empirical bivariate marginal tail dependence function as,

$$\lambda_{i,j}^{emp}(x_i, x_j; k) := \frac{1}{k} \sum_{h=1}^n \mathbf{I} \left\{ 1 - F_i(x_{h,i}) \leq \frac{k}{n} x_i, 1 - F_j(x_{h,j}) \leq \frac{k}{n} x_j \right\} \quad (3.11)$$

Since only tail events are considered $k(n) \rightarrow \infty$ and $\frac{k}{n} \rightarrow 0$ as $n \rightarrow \infty$. ■

There exist many more empirical tail dependence estimates for more detail see Schmidt and Stadtmüller (2006).

The next section deals with how, instead of empirical measures, functions known as copulas can be used to describe the dependence between random variables.

3.2 COPULA THEORY

A crucial element of multivariate statistics is describing the probability distribution of some random vector $(X_1, \dots, X_d)'$. The probability distribution function of $(X_1, \dots, X_d)'$ is characterised as follows:

$$F(x_1, \dots, x_d) := P(X_1 \leq x_1, \dots, X_d \leq x_d), \quad x_1, \dots, x_d \in \mathbb{R} \quad (3.12)$$

3.2. COPULA THEORY

However, barring some well known multivariate distributions, expressing $F(x_1, \dots, x_d)$ in (3.12) mathematically can be complex and be very specific to data with particular attributes. A more convenient approach may be to consider the probability distribution or marginal distribution $F_i(x) := \mathbb{P}(X_i < x)$ of each of the random variables $X_i, i = 1, 2, \dots, d$ individually. Then once each random variable can be described by some known marginal distribution, combining these marginals to a multivariate distribution can then be pursued. Thus in addition to the description of the marginal distribution for each variable, a function is required to combine these marginals in a mathematically sound manner. Furthermore, such a function would not only combine these marginal distributions but as a consequence also fully describe the dependence between the random variables. A function that constructs a multivariate distribution from underlying marginal distributions is known as a copula. The word copula originates from Latin to mean “a link, tie or bond” (Nelsen, 2007).

Copulas was first described by Sklar (1959) as a function that joins one-dimensional probability distributions to form multivariate probability distributions. Further, he detailed a theorem which is now considered to be the fundamental theorem in the field of copulas and as such it is termed Sklar’s Theorem.

Theorem 3.1 (Sklar’s Theorem). *A function $F : \mathbb{R}^d \rightarrow [0, 1]$ is the distribution function of some random vector $(X_1, \dots, X_d)'$ if and only if there is a copula $C : [0, 1]^d \rightarrow [0, 1]$ and univariate distribution functions $F_1, \dots, F_d : \mathbb{R} \rightarrow [0, 1]$ such that*

$$C(F_1(x), \dots, F_d(x)) = F(x_1, \dots, x_d), \quad x_1, \dots, x_d \in \mathbb{R} \quad (3.13)$$

If marginals F_1, \dots, F_d are continuous, then C is unique. Conversely, if C is a copula and F_1, \dots, F_d are univariate distribution functions, then the function F defined in (3.13) is a joint distribution function with margins F_1, \dots, F_d . ■

Similarly, Sklar’s Theorem can be applied to construct multivariate survival functions. A survival function of a random vector $(X_1, \dots, X_d)'$ is defined as

$$\bar{F}(x_1, \dots, x_d) = P(X_1 > x_1, \dots, X_d > x_d), \quad x_1, \dots, x_d \in \mathbb{R} \quad (3.14)$$

Each random variable is specified by its marginal survival function $\bar{F}_j(x) := P(X_j > x) = 1 - F_j(x), x \in \mathbb{R}, j = 1, 2, \dots, d$. To construct a multivariate survival function from

3.2. COPULA THEORY

the marginal survival functions a survival copula \hat{C} is employed. Furthermore, Sklar's Theorem can be reformulated for survival copulas as follows.

Theorem 3.2 (Sklar's Theorem for Survival Functions). *A function $\bar{F} : \mathbb{R}^d \rightarrow [0, 1]$ is the distribution function of some random vector $(X_1, \dots, X_d)^t$ if and only if there is a survival copula $\hat{C} : [0, 1]^d \rightarrow [0, 1]$ and univariate survival functions $\bar{F}_1, \dots, \bar{F}_d : \mathbb{R} \rightarrow [0, 1]$ such that*

$$\hat{C}(F_1(x), \dots, F_d(x)) = \bar{F}(x_1, \dots, x_d), \quad x_1, \dots, x_d \in \mathbb{R} \quad (3.15)$$

If marginals $\bar{F}_1, \dots, \bar{F}_d$ are continuous, then \hat{C} is unique. Conversely, if \hat{C} is a survival copula and $\bar{F}_1, \dots, \bar{F}_d$ are univariate survival functions, then the function \bar{F} defined in (3.15) is a joint survival function with margins $\bar{F}_1, \dots, \bar{F}_d$. ■

Since the publication of Sklar's Theorem, copulas have been applied over a wide variety of fields. Copulas was especially applied in the field of finance and according to Salmon (2012) is partly to blame for the 2008/9 credit crises. A more thorough mathematical examination of copulas can be found in Nelsen (2007).

As stated earlier, since a copula combines the marginals of all random variables it, as a consequence, also fully describes the dependence between all random variables. Therefore, copulas can be used not only to construct multivariate distributions, but also to analyse the dependence in a multivariate dataset. Given that the topic of this study is to apply alternative dependence measures to biplots, copulas will be studied for its ability to characterise dependence. If a copula fully characterises dependence then it is obvious that a copula should, to an extent, be associated with alternative dependence measures such as Kendall's tau and Spearman's rho, which is described by Nelsen (2007) in the following theorems.

Theorem 3.3 (Kendall's tau Copula Specification). *Let X and Y be continuous random variables whose copula is C . Further, let F_X and F_Y be the respective marginal distribution functions and define a random vector of uniform variables by $(U, V) := [F_X(X), F_Y(Y)]$. Then Kendall's tau for X and Y is given by*

$$\tau_{X,Y} = \tau_C := 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1 = 4\mathbb{E}[C(U, V)] - 1 \quad (3.16)$$

■

3.2. COPULA THEORY

Theorem 3.4 (Spearman's rho Copula Specification). *Let X and Y be continuous random variables whose copula is C . Further, let F_X and F_Y be the respective marginal distribution functions and define a random vector of uniform variables by $(U, V) := [F_X(X), F_Y(Y)]$. Then Kendall's tau for X and Y is given by*

$$\rho_{X,Y} = \rho_C := 12 \int_0^1 \int_0^1 C(u, v) du dv - 3 \quad (3.17)$$

■

The relationships described in Theorems 3.3 and 3.4 are useful when fitting a copula to a dataset. Given that Kendall's tau and Spearman's rho are easy to estimate from observations, the relationships above can be used to estimate the parameters for a parametric copula.

Finally, it is further shown by Nelsen (2007) that the tail dependence coefficient depends solely on the underlying copula of a multivariate distribution. This is conveyed by the following theorem.

Theorem 3.5. *Let X and Y be continuous random variables whose copula is C . Further, let F_X and F_Y be the respective marginal distribution functions and define a random vector of uniform variables by $(U, V) := [F_X(X), F_Y(Y)]$. Then the upper- and lower-tail dependence coefficients for X and Y is given by*

$$\lambda_{upper}(U, V) := \lim_{t \rightarrow 1} \frac{C(t, t) - 2t + 1}{1 - t} \quad (3.18)$$

and

$$\lambda_{lower}(U, V) := \lim_{t \rightarrow 0} \frac{C(t, t)}{t} \quad (3.19)$$

■

Given that estimating the tail dependence coefficients is challenging since it entails estimating a property in a limit from finite observations, it is simpler to first fit a copula to the data and then use the relationship in Theorem 3.5 to determine the upper- and lower-tail dependence coefficients.

In the subsequent section, some popular copula families and their properties are examined.

3.2. COPULA THEORY

3.2.1 Elliptical copulas

One of the central themes of this study is to find a method to characterise extreme dependence. To achieve this, properties from an elliptical copula will be applied. Therefore, this section is devoted to presenting the necessary background on elliptical copulas.

Before elliptical copulas can be discussed it is first necessary to ask: *What is an elliptical distribution?* The family of elliptical distributions consists of distributions that generalise the multivariate normal distribution, for example the multivariate Student-t distribution. An elliptical distribution is constructed by combining elliptical marginal distributions through an elliptical copula. More generally, an elliptical distribution can be obtained through a linear transformation of a spherical distribution. A spherical distribution is defined as follows:

Definition 3.3 (Spherical distribution). *Suppose that $\mathbf{O} : d \times d$ is an orthogonal matrix such that $\mathbf{O}'\mathbf{O} = I$. Then a d -dimensional random vector $\mathbf{X} = (X_1, \dots, X_d)'$ has a spherical distribution if for every matrix \mathbf{O} one has*

$$\mathbf{O}\mathbf{X} \stackrel{d}{=} \mathbf{X} \quad (3.20)$$

Equivalently, there exists a random variable $R \geq 0$ and, independently, a d -dimensional random vector \mathbf{S} with a uniform distribution on an unit sphere, such that

$$\mathbf{X} \stackrel{d}{=} R\mathbf{S}. \quad (3.21)$$

■

Furthermore, a spherical distribution is fully described through a function $\phi : [0, \infty) \rightarrow \mathbb{R}$, referred to as a characteristic or generator function of \mathbf{X} . Hence, the distribution of a d -dimensional spherical random variable is denoted as $\mathbf{X} \sim S_d(\phi)$. Since an elliptical distribution is merely a linear transformation of a spherical distribution the following theorem holds.

Theorem 3.6. *A d -dimensional random vector \mathbf{Z} is said to have an elliptical distribution, if and only if, there exists a non-negative random variable R independently of \mathbf{S} , a*

3.2. COPULA THEORY

d -dimensional random vector on a unit hypersphere, and a $d \times k$ matrix A with $AA' = \Sigma$, such that

$$\mathbf{Z} \stackrel{d}{=} \boldsymbol{\mu} + A' \mathbf{X} \stackrel{d}{=} \boldsymbol{\mu} + RA'S \quad (3.22)$$

with $\mathbf{X} \sim S_d(\phi)$. ■

From Theorem 3.6 the definition of an elliptical distribution follows directly.

Definition 3.4 (Elliptical distribution). *If \mathbf{Z} is a d -dimensional random vector and, for some vector $\boldsymbol{\mu} \in \mathbb{R}^d$, some $d \times d$ non-negative definite symmetric matrix Σ , and some function $\phi: [0, \infty) \rightarrow \mathbb{R}$, the characteristic function $\varphi_{\mathbf{Z}-\boldsymbol{\mu}}$ of $\mathbf{Z}-\boldsymbol{\mu}$ is of the form $\varphi_{\mathbf{Z}-\boldsymbol{\mu}}(\mathbf{t}) = \phi(\mathbf{t}'\Sigma\mathbf{t})$, we say that \mathbf{Z} has an elliptical distribution with parameters $\boldsymbol{\mu}$, Σ and ϕ , and we write $\mathbf{Z} \sim \mathcal{E}_d(\boldsymbol{\mu}, \Sigma, \phi)$.* ■

Further, Klüppelberg and Kuhn (2009) proves that if $\mathbf{Z} \sim \mathcal{E}_d(\boldsymbol{\mu}, \Sigma, \phi)$ is transformed as,

$$\mathbf{Z}^* := \text{diag}(\sigma_{11}, \dots, \sigma_{dd})^{\frac{1}{2}} \mathbf{Z} \quad (3.23)$$

then $\mathbf{Z}^* \sim \mathcal{E}_d(\boldsymbol{\mu}, \mathcal{R}, \phi)$ with $\mathcal{R} := (\sigma_{ij} / \sqrt{\sigma_{ii}\sigma_{jj}})_{i \leq j, j \leq d}$ the correlation matrix of \mathbf{Z} . Hence under a suitable transformation an elliptical copula can be specified via its correlation matrix and characteristic function ϕ .

The correlation matrix from an elliptical distribution has a useful relationship with Kendall's tau, which is presented in Theorem 3.7.

Theorem 3.7. *Let $\mathbf{Z} \sim \mathcal{E}_d(\boldsymbol{\mu}, \mathcal{R}, \phi)$, where for $i, j \in 1, 2, \dots, d$, Z_i and Z_j are continuous with correlation coefficient ρ_{ij} . Then,*

$$\tau(Z_i, Z_j) = \frac{2}{\pi} \arcsin \rho_{ij}. \quad (3.24)$$
■

This relationship can be used to find for the correlation coefficient ρ_{ij} by estimating $\tau(Z_i, Z_j)$ and inverting (3.24).

Before the extremal dependence properties of an elliptical distribution can be examined, it is first necessary to define the concept of a regularly varying random variable. A regularly varying random variable is defined as follows:

3.2. COPULA THEORY

Definition 3.5 (Regular variation). *A random variable R is said to be regularly varying with index $\alpha > 0$ if for all $x > 0$,*

$$\lim_{t \rightarrow \infty} \frac{P(R > tx)}{P(R > t)} = x^{-\alpha} \quad (3.25)$$

A regularly varying random variable is merely a random variable whose tail behaviour is similar to that of a power law function. According to Hult and Lindskog (2002) the concept of regular variation is connected to the tail dependence function of an elliptical distribution in the following theorem.

Theorem 3.8. *Let $\mathbf{Z} \stackrel{d}{=} \boldsymbol{\mu} + R\mathbf{A}'\mathbf{S} \sim \mathcal{E}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \phi)$ with $\Sigma_{ii} > 0$ for $i = 1, 2, \dots, d$, $|\rho_{ij}| < 1$ for all $i \neq j$, and where $\boldsymbol{\mu}$, R , \mathbf{A} and \mathbf{S} are as in Theorem 3.6. Then the following statements are equivalent:*

- i. R is regularly varying with index $\alpha > 0$.
- ii. \mathbf{Z} is regularly varying with index $\alpha > 0$.
- iii. For all $i \neq j$, (Z_i, Z_j) has tail dependence coefficient

$$\lambda_U(Z_i, Z_j) = \lambda_\ell(Z_i, Z_j) = \frac{\int_{(\pi/2 - \arcsin \rho_{ij})/2}^{\pi/2} \cos^\alpha(t) dt}{\int_0^{\pi/2} \cos^\alpha(t) dt}. \quad (3.26)$$

Hult and Lindskog (2002) concludes from Theorem 3.8 that an elliptically distributed random vector \mathbf{Z} will only have tail dependence if the random variable R in $\mathbf{Z} \stackrel{d}{=} \boldsymbol{\mu} + R\mathbf{A}'\mathbf{S}$ is regularly varying. Moreover, the correlation coefficient ρ_{ij} only affects the magnitude of the tail dependence. This consequently implies that a bivariate observation (Z_i, Z_j) can have a large tail dependence coefficient even if its correlation coefficient is zero.

A more generalised approach to characterise the dependence for an elliptical distribution is through an elliptical copula which is defined as follows.

Definition 3.6 (Elliptical copula). *An elliptical copula is defined as the copula related to an elliptical distribution F , and is obtained using Sklar's theorem as,*

$$C(u_1, \dots, u_d) = F[F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)], \quad u_1, \dots, u_d \in [0, 1]^d \quad (3.27)$$

where F_k^{-1} , $k = 1, 2, \dots, d$ are univariate quantile functions. ■

3.2. COPULA THEORY

Two of the most well known elliptical distributions is the multivariate- normal and Student-t distributions, both having corresponding elliptical copulas defined as follows:

Definition 3.7 (Gaussian copula). *The Gaussian copula $C_{\mathcal{R}}^{Guass}$ is the copula of $\mathbf{X} \sim N_d(\mathbf{0}, \mathcal{R})$, where \mathcal{R} is the correlation matrix of \mathbf{X} . The analytical form is given by*

$$C_{\mathcal{R}}^{Guass} := \Phi_{\mathcal{R}}[\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d)] \quad (3.28)$$

where $(u_1, \dots, u_d) \in [0, 1]^d$, $\Phi_{\mathcal{R}}$ is the joint distribution function of \mathbf{X} , and Φ^{-1} is the quantile function of the standard normal distribution. ■

Definition 3.8 (t-Copula). *The t-Copula $C_{\mathcal{R}, \nu}^t$ is the copula of $\mathbf{X} \sim t_d(\mathbf{0}, \mathcal{R}, \nu)$, where \mathcal{R} is the correlation matrix of \mathbf{X} . The analytical form is given by*

$$C_{\mathcal{R}, \nu}^t := t_{\nu, \mathcal{R}}[t_{\nu}^{-1}(u_1), \dots, t_{\nu}^{-1}(u_d)] \quad (3.29)$$

where $(u_1, \dots, u_d) \in [0, 1]^d$, $t_{\nu, \mathcal{R}}$ is the joint distribution function of \mathbf{X} , and t_{ν}^{-1} is the quantile function of the Student-t distribution with ν degrees of freedom. ■

Further, the upper and lower tail dependence of the Gaussian copula is zero, i.e the Gaussian copula is tail independent. However, for the t-Copula the upper and lower tail dependence is identical and is expressed as

$$\lambda_U(X_i, X_j) = \lambda_L(X_i, X_j) = 2t_{\nu+1} \left(-\sqrt{\frac{(\nu+1)(1-\rho_{ij})}{1+\rho_{ij}}} \right) \quad (3.30)$$

with ν the degrees of freedom.

Algorithms 3.1 and 3.2, as given by Embrechts *et al.* (2002), provides the steps to generate multivariate samples using the Gaussian and t-Copula, respectively.

Algorithm 3.1 (Simulating n multivariate samples from the Gaussian copula). *In order to generate n observations from a d -variate distribution with marginal distributions F_1, F_2, \dots, F_d and quantile functions $F_1^{-1}, F_2^{-1}, \dots, F_d^{-1}$ that is specified and known. Additionally, suppose that dependence structure is characterised by a Gaussian copula with correlation matrix \mathcal{R} . Then perform the following steps:*

3.2. COPULA THEORY

1. Compute the Cholesky decomposition of \mathcal{R} by computing the lower triangular matrix $\mathbf{L}: d \times d$, with $\mathbf{L}\mathbf{L}' = \mathcal{R}$.
2. Simulate a vector of d independent standard normal random variables $\mathbf{z} = (z_1, \dots, z_d) \sim N_d(\mathbf{0}, \mathbf{I})$.
3. Compute $\mathbf{y} = (y_1, \dots, y_d) := \mathbf{L}\mathbf{z} \sim N_d(\mathbf{0}, \mathbf{L}\mathbf{L}') = N_d(\mathbf{0}, \mathcal{R})$.
4. Return the vector $\mathbf{u} = (u_1, \dots, u_d) = (\Phi(y_1), \dots, \Phi(y_d))$.
5. Use the Probability Inverse Transformation to generate the sample as $(x_1, \dots, x_d) = [F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)]$.
6. Repeat steps 1 to 6 n times to generate n samples.

Algorithm 3.2 (Simulating n multivariate samples from the t -Copula). *In order to generate n observations from a d -variate distribution with marginal distributions F_1, F_2, \dots, F_d and quantile functions $F_1^{-1}, F_2^{-1}, \dots, F_d^{-1}$ that is specified and known. Additionally, suppose that dependence structure is characterised by a t -Copula with correlation matrix \mathcal{R} and ν degrees of freedom. Then perform the following steps:*

1. Compute the Cholesky decomposition of \mathcal{R} by computing the lower triangular matrix $\mathbf{L}: d \times d$, with $\mathbf{L}\mathbf{L}' = \mathcal{R}$.
2. Simulate a vector of d independent standard normal random variables $\mathbf{z} = (z_1, \dots, z_d) \sim N_d(\mathbf{0}, \mathbf{I})$.
3. Simulate random variable W from $\chi^2(\nu)$.
4. Compute $\mathbf{y} = (y_1, \dots, y_d) := \sqrt{\frac{\nu}{W}}\mathbf{L}\mathbf{z}$.
5. Return the vector $\mathbf{u} = (u_1, \dots, u_d) = (t_\nu(y_1), \dots, t_\nu(y_d))$.
6. Use the Probability Inverse Transformation to generate the sample as $(x_1, \dots, x_d) = [F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)]$.
7. Repeat steps 1 to 6 n times to generate n samples.

3.2. COPULA THEORY

The differences in the Gaussian and t-Copula is illustrated in Figure 3.1, with 10,000 simulations from a: (A) bivariate normal distribution, (B) bivariate t_4 distribution, (C) bivariate random variables with normal marginals and a t_4 -Copula, and (D) bivariate random variables with t_4 marginals and a Gaussian copula. The corresponding Gaussian and t-Copula with parameter $\rho = 0.9$ is presented in Figure 3.2(A) and 3.2(B), respectively.

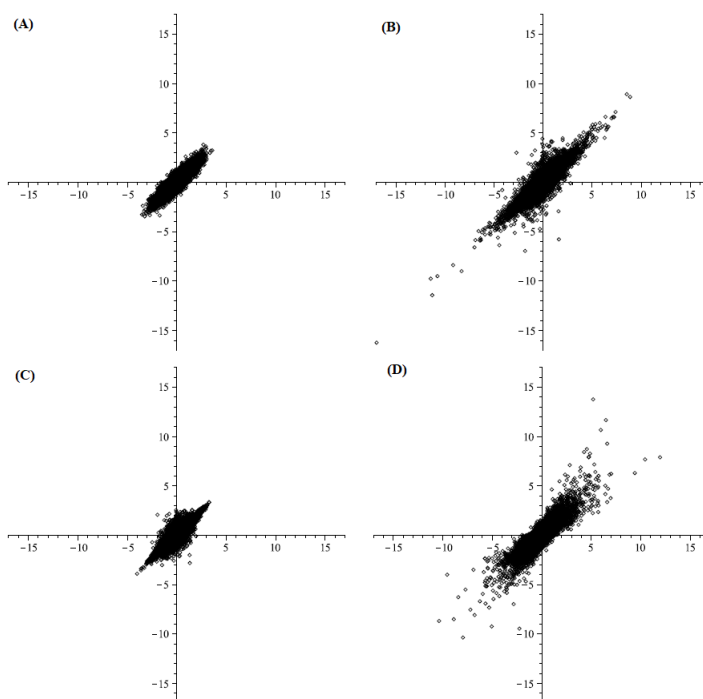


Figure 3.1

Simulation of 10000 bivariate normal and t_4 distributed random variables

Source: Klüppelberg and Stelzer (2014)

3.2. COPULA THEORY

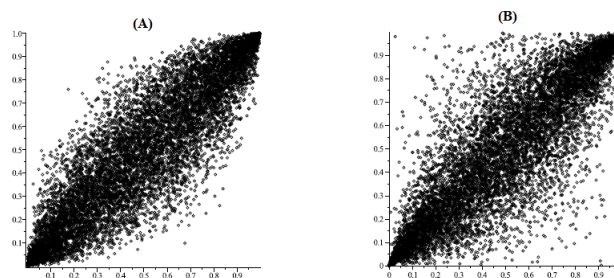


Figure 3.2

Scatter plot of univariate marginals corresponding to the Gaussian and t_4 -Copula

Source: Klüppelberg and Stelzer (2014)

When comparing Figure 3.1(A) and (B), both with normal marginals, it is clear that the t_4 -Copula in (B) produces more observations in the upper-right and lower-left quadrants as compared to the Gaussian copula in (A). Furthermore, considering the t_4 marginals in Figure 3.1(C) and (D) with t-Copula and Gaussian copula, respectively. The first noticeable difference is that the t_4 marginals have heavier tails. Additionally, the observation with t_4 marginals and Gaussian copula in (D) is more dispersed in the lower-right and upper-left quadrants. The difference in (B) and (D) illustrates the tail independence of the Gaussian copula. Finally, Figure 3.2(A) and (B) presents scatter plots of marginals under the Gaussian and t-Copula. As can be seen, the Gaussian copula (A) has fewer observations in the upper right and lower left corners as compared to the t-Copula in (B).

3.2.2 Archimedean copulas

The popularity of multivariate elliptical distribution has directly lead to elliptical copulas, such as the Gaussian copula, being widely used in finance. Unfortunately, the Gaussian copula is not always well suited for modelling financial data, as was seen during the 2008/09 financial crises when it was found that the Gaussian copula was ill-suited to model CDO losses (see Salmon, 2012).

Mai and Scherer (2014) states that elliptical copulas have the drawback of not only having a complicated algebraic expression but it is also subject to a great level of symme-

3.2. COPULA THEORY

try. The family of Archimedean copulas addresses these drawbacks by having a convenient algebraic form and having the ability to account for asymmetries. To define Archimedean copulas it is first necessary to describe a mathematical result known as the Laplace Transformation (LT). The LT of a non-negative random variable X with distribution function F_X is defined as:

$$\varphi_X(s) := \mathbb{E}[e^{-sX}] = \int_0^{\infty} e^{-sx} dF_X(x), \quad s > 0 \quad (3.31)$$

The intuition behind Archimedean copulas is explained by Matthias and Jan-frederik (2017) as follows.

Consider a vector of *i.i.d.* Exponential random variables $E_1, E_2, \dots, E_d \sim \text{Exp}(1)$. Furthermore, independent of this let M be a positive random variable. Define a vector of random variables by

$$(X_1, \dots, X_d)' := \left(\frac{E_1}{M}, \dots, \frac{E_d}{M} \right)' \quad (3.32)$$

Then it can be shown that the survival function of the k^{th} , $k = 1, 2, \dots, d$, random variable can be expressed as

$$\bar{F}_k(x) = \mathbb{E}[e^{-xM}] = \int_0^{\infty} e^{-mx} dF_M(m) =: \varphi(x), \quad x \geq 0 \quad (3.33)$$

where φ is the LT of M . Then as long as M is not deterministic, the random variables $(X_1, \dots, X_d)'$ are dependent since each is affected by M . For this reason, it can be proven that the survival copula of the random variables $(X_1, \dots, X_d)'$ is parameterised by the LT of M , and is given by

$$\hat{C}_\varphi(u_1, \dots, u_d) := \varphi(\varphi^{-1}(u_1) + \dots + \varphi^{-1}(u_d)) \quad (3.34)$$

where $u_1, \dots, u_d \in [0, 1]$ and φ is the LT of M . Thus the survival function for $(X_1, \dots, X_d)'$ can be expressed as,

$$P(X_1 > x_1, \dots, X_d > x_d) = \hat{C}_\varphi(\varphi(x_1), \dots, \varphi(x_d)) \quad (3.35)$$

where $x_1, \dots, x_d > 0$.

Copulas derived in this manner are termed Archimedean copulas and has the functional form as presented in (3.34). Hence, the definition of Archimedean copulas is as follows:

3.2. COPULA THEORY

Definition 3.9 (Archimedean copula). *A d -dimensional copula is called Archimedean if it admits to the functional form*

$$C_\varphi(u_1, \dots, u_d) := \varphi(\varphi^{-1}(u_1) + \dots + \varphi^{-1}(u_d)) \tag{3.36}$$

where $u_1, \dots, u_d \in [0, 1]$ and the function $\varphi : [0, \infty) \rightarrow [0, 1]$ with inverse φ^{-1} is termed the generator of C_φ . ■

Each copula within the family of Archimedean copulas possesses a distinct generator function parameterised by θ that determines the level of dependence for the copula. Three of the most well studied Archimedean copulas, each with their own distinct properties, are the Gumbel (C_θ^{Gu}), Clayton (C_θ^{Cl}), and Frank (C_θ^{Fr}) copulas. These copulas with their corresponding properties are presented in Table 3.1.

Table 3.1
Popular families of Archimedean copulas

Copula	Parameter (θ)	Generator ($\varphi_\theta(x)$)	Independence	Comonotonicity	Kendall's tau (τ_θ)
Gumbel (C_θ^{Gu})	$\theta \in [1, \infty)$	$(1+x)^{-\frac{1}{\theta}}$	$\theta \rightarrow 1$	$\theta \rightarrow \infty$	$\frac{\theta-1}{\theta}$
Clayton (C_θ^{Cl})	$\theta \in (0, \infty)$	$\exp(-x^{\frac{1}{\theta}})$	$\theta \rightarrow 0$	$\theta \rightarrow \infty$	$\frac{\theta}{2+\theta}$
Frank (C_θ^{Fr})	$\theta \in (0, \infty)$	$-\frac{1}{\theta} \ln[e^{-x}(e^{-\theta} - 1) + 1]$	$\theta \rightarrow 0$	$\theta \rightarrow \infty$	$1 + 4 \frac{\int_0^\theta t(e^t - \theta)^{-1} dt - 1}{\theta}$

Algorithm 3.3 can be used to simulate observations from an Archimedean copula.

Algorithm 3.3 (Simulating from Archimedean copulas). *In order to generate n observations from a d -variate distribution with marginal distributions F_1, F_2, \dots, F_d and quantile functions $F_1^{-1}, F_2^{-1}, \dots, F_d^{-1}$ that is specified and known. Additionally, suppose that the dependence structure is characterised by an Archimedean copula with parameter θ . Then perform the following steps:*

1. Specify parameter θ such that:
 - (a) $\theta > 1$ for the Gumbel Copula.
 - (b) $\theta > 0$ for the Clayton Copula.

3.2. COPULA THEORY

(c) $\theta > 0$ for the Frank Copula.

2. Generate a

(a) Stable variate $V \sim St(1/\theta, 1, \gamma, 0)$, where $\gamma = [\cos(\pi/\theta)]^\theta$ with the Laplace-Transform of its distribution function given by $\varphi_\theta(t) = \exp(-t^{\frac{1}{\theta}})$.

(b) Gamma variate $V \sim Ga(1/\theta, 1)$ with $\theta > 0$, with the Laplace-Transform of its distribution function is given by $\varphi_\theta(t) = (1 + t)^{-1/\theta}$.

(c) Discrete variate V with probability mass function $p(k) = \mathbb{P}(V = k) = (1 - \exp(-\theta))^k / (k\theta)$ for $k = 1, 2, \dots$ and $\theta > 0$ with $\varphi_\theta(t) = -\frac{1}{\theta} \ln [e^{-x}(e^{-\theta} - 1) + 1]$.

3. Generate independent Uniform variables $Y_1, \dots, Y_d \sim U(0, 1)$.

4. Return $(u_1, \dots, u_d) = \left[\varphi_\theta^{-1}\left(-\frac{\ln(y_1)}{V}\right), \dots, \varphi_\theta^{-1}\left(-\frac{\ln(y_d)}{V}\right) \right]$.

5. Use the Probability Inverse Transform to generate the sample $(x_1, \dots, x_d) = [F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)]$.

6. Repeat steps 1 to 5 n times to generate n samples.

■

The Gumbel copula (C_θ^{Gu}) has no lower tail dependence, only upper tail dependence of $2 - 2^{\frac{1}{\theta}}$. Furthermore, C_θ^{Gu} is the only Archimedean copula that is also an extreme value copula. A copula C is said to be an extreme value copula if it satisfies the property

$$C(u_1, \dots, u_d)^t = C(u_1^t, \dots, u_d^t), \quad \forall t \geq 0 \quad \text{and} \quad u_1, \dots, u_d \in [0, 1] \quad (3.37)$$

Such copulas play an important role in extreme value theory but are beyond the scope of this study.

Simulations for a bivariate Gumbel copula is illustrated in Figure 3.3 where it can be seen that the dependence increases as θ increases. Additionally, it can be seen that observations appear to be more correlated in the upper right corner of the scatter plot, illustrating the upper tail dependence present in the Gumbel copula.

3.2. COPULA THEORY

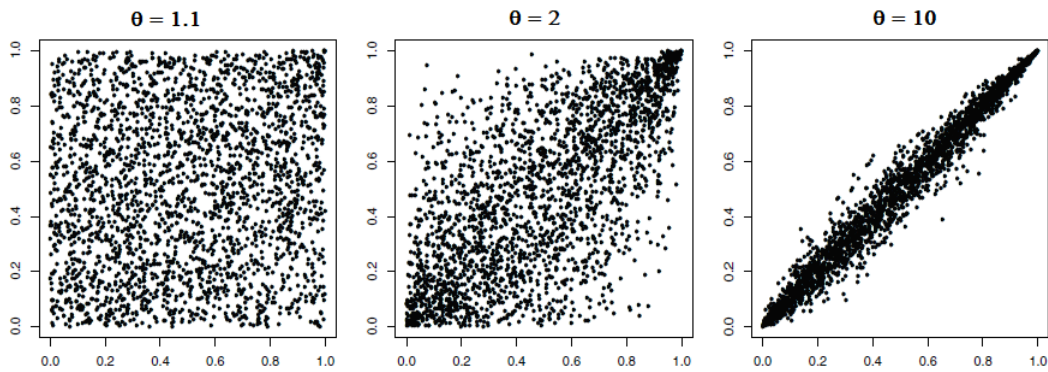


Figure 3.3

Scatter plots of univariate marginals corresponding to a Gumbel copula

Source: Matthias and Jan-frederik (2017)

The Clayton copula (C_{θ}^{Cl}), in contrast to the Gumbel copula, has no upper tail dependence, but a lower tail dependence of $2^{-\frac{1}{\theta}}$. The bivariate Clayton copula is illustrated in Figure 3.4 and, in a similar manner, dependence increases as θ becomes larger. Furthermore, since the Clayton copula is subject to lower tail dependence, one observes greater dependence in the lower left corner of the scatter plot.

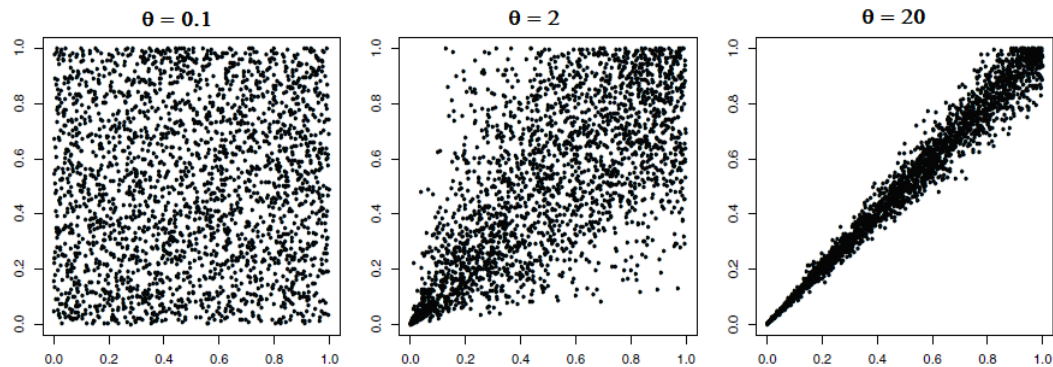


Figure 3.4

Scatter plots of univariate marginals corresponding to a Clayton copula

Source: Matthias and Jan-frederik (2017)

Finally, the Frank copula (C_{θ}^{Fr}) is tail independent and has the useful property in the bivariate case whereby its copula for the distribution and survival function are the same.

3.3. MULTIVARIATE EXTREME DEPENDENCE ANALYSIS

The scatter plot for the Frank copula is presented in Figure 3.5 whereby it is clear that no tail dependence is present.

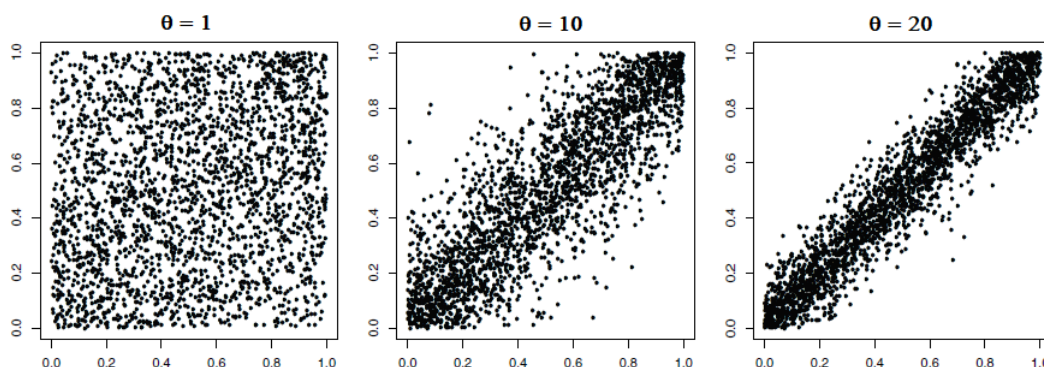


Figure 3.5

Scatter plots of univariate marginals corresponding to a Frank copula

Source: Matthias and Jan-frederik (2017)

3.3 MULTIVARIATE EXTREME DEPENDENCE ANALYSIS

It is well established in the financial world that dependence between asset returns drastically changes during volatile markets and in extreme economic circumstances (see Longin and Solnik, 2001). This implies that special attention should be given to the dependence of extreme observations, given that by Theorem 3.5, a relationship exists between copulas and extremal dependence measures such as the tail dependence coefficient. This suggests that copulas can be used as a tool to analyse extreme dependence. Therefore, if a copula structure is assumed to determine the dependence of extreme observations, inferences can be made about the overall strength of such an extreme dependence. More specifically, by assuming that extreme observations are obtained from an elliptical distribution with its dependence described by an elliptical copula, a correlation matrix for extreme observations can be determined. Such an approach was studied by Haug *et al.* (2015) with the goal of deriving a correlation matrix from extreme observations.

3.3. MULTIVARIATE EXTREME DEPENDENCE ANALYSIS

This is done by making the assumption that a random vector $\mathbf{X} = (X_1, \dots, X_d)'$ is generated from an elliptical distribution with regularly varying generating function ϕ with index ν . Further, Klüppelberg *et al.* (2007) shows that the bivariate marginal tail dependence coefficient for a regularly varying elliptical distribution can be determined using the following theorem.

Theorem 3.9. *Let $\mathbf{X} = (X_1, \dots, X_d)'$ be a random vector with elliptical copula $C_d(\mathcal{R}, \phi)$.*

Then the bivariate marginal tail dependence function of \mathbf{X} is given by

$$\begin{aligned} \lambda(x_i, x_j, \nu, \rho_{ij}) &:= \left[x_i \int_{\phi_{ij}[(x_i/x_j)^{1/\nu}]}^{\pi/2} [\cos(\theta)]^\nu d\theta + x_j \int_{\phi_{ij}[(x_j/x_i)^{-1/\nu}]}^{\pi/2} [\cos(\theta)]^\nu d\theta \right] \left[\int_{\pi/2}^{\pi/2} [\cos(\theta)]^\nu d\theta \right]^{-1} \\ &= x_i \left[1 - t_{\nu+1} \left(\frac{(x_i/x_j)^{1/\nu} - \rho_{ij} \sqrt{\nu+1}}{\sqrt{1-\rho_{ij}^2}} \right) \right] + x_j \left[1 - t_{\nu+1} \left(\frac{(x_j/x_i)^{1/\nu} - \rho_{ij} \sqrt{\nu+1}}{\sqrt{1-\rho_{ij}^2}} \right) \right] \end{aligned} \quad (3.38)$$

where x_i and x_j are the respective i^{th} and j^{th} components of \mathbf{X} . Moreover, $\phi_{ij}(t) := \arctan \left[(t - \rho_{ij}) / \sqrt{1 - \rho_{ij}^2} \right]$ and $t_{\nu+1}$ denotes the t -distribution function with $\nu + 1$ degrees of freedom.

The above theorem is a generalisation of the tail dependence coefficient given in Theorem 3.8. This is because the tail dependence coefficient in Theorem 3.8 only considers points at $[1, 1]$ whereas the above considers points $[\sqrt{2} \cos(\theta), \sqrt{2} \sin(\theta)]$ for $\theta \in [0, \pi/2]$, which includes the point $(1, 1)$.

Now Haug *et al.* (2015) gives an algorithm to estimate $\rho_{i,j}$ for each pair of variables (x_i, x_j) , calibrated from extreme observations by using the relationship provided in Theorem 3.9. The tail dependence function in (3.38) is a function of two parameters ν and ρ_{ij} and can hence be denoted as the function $\lambda(x_i, x_j; \nu; \rho_{ij})$. Now, estimating two parameters ν and ρ_{ij} from one function is not straight forward and requires a two step procedure. Since the end goal is to find p_{ij} the function in (3.38) needs to be inverted with respect to $\rho_{ij} := \rho_{ij}(x_i, x_j; \nu; \lambda)$ implying that estimates for ν and λ must be found.

Finding an estimator for λ is straightforward by using the empirical bivariate marginal tail dependence coefficient given in definition 3.2 as

$$\hat{\lambda}_{i,j}^{emp}(x_i, x_j; k) := \frac{1}{k} \sum_{h=1}^n \mathbf{I} \left\{ 1 - F_i(x_{h,i}) \leq \frac{k}{n} x_i, 1 - F_j(x_{h,j}) \leq \frac{k}{n} x_j \right\} \quad (3.39)$$

3.4. SUMMARY

with k the number of tail samples. Hence the estimator for the bivariate marginal tail dependence coefficient is taken to be $\hat{\lambda}_{i,j}^{emp}$.

Then in order to estimate the second parameter ν , (3.38) is inverted as a function of ν denoted as $\nu(\hat{\lambda}_{i,j}^{emp}; \rho_{ij})$ with $\hat{\lambda}_{i,j}^{emp}$ substituted as an estimate for the bivariate marginal tail dependence coefficient. Note that, in order to estimate ν requires an estimate of ρ_{ij} , in spite of it being the parameter that is suppose to be determined at the end. To overcome this Haug *et al.* (2015) suggest using an initial estimator of ρ_{ij} that is only used to determine ν . This initial estimator of ρ_{ij} is found be estimating Kendall's tau $\hat{\tau}_{ij}$ for random variables and using the relationship in Theorem 3.7 to solve for ρ_{ij} as,

$$\hat{\rho}_{ij}^{\tau} := \sin\left(\frac{\pi}{2}\hat{\tau}_{ij}\right) \quad (3.40)$$

As a result, ν can be estimated using the above estimator as $\hat{\nu} := \nu(\hat{\lambda}_{i,j}^{emp}; \hat{\rho}_{ij}^{\tau})$. Finally, both estimators $\hat{\nu}$ and $\hat{\lambda}_{i,j}^{emp}$ are now employed to estimate the correlation coefficient for random variables X_i and X_j as $\hat{\rho}_{ij} = \rho_{ij}(x_i, x_j; \hat{\nu}; \hat{\lambda}_{i,j}^{emp})$. As a result, $\hat{\rho}_{ij}$ is the extremal correlation coefficient for X_i, X_j determined from extreme observations. This is then repeated to find extremal correlation coefficient for all combinations of i, j which are then combined into an extremal correlation matrix denoted $\hat{\mathcal{R}}^{extreme}$.

3.4 SUMMARY

In this chapter, an in-depth look was taken at the concept and measurement of dependence. Initially, some background on traditional dependence measures was provided such as Pearson's correlation, Kendall's tau and Spearman's rho. Subsequently, copulas were introduced as a functional approach to characterise dependence. This was followed by an examination of various families of copulas which will be applied in later chapters to simulate multivariate data with various properties. Finally, a methodology to analyse multivariate extreme dependence was presented. This methodology yields a correlation matrix calibrated from extremes that will be applied in the next chapter to construct PCA biplots more suited to extremes.

CHAPTER 4

METHODOLOGY

In Chapter 2 background regarding the development and application of PCA biplots were discussed. Then in Chapter 3, concepts related to dependence and the measurement of dependence was discussed. Further, a methodology to determine a correlation matrix calibrated from extreme observations was presented in section 3.2. With these topics in mind the purpose of this chapter is to combine key ideas from Chapter 3 to the PCA biplot methodology discussed in Chapter 2. The intention of this new PCA biplot methodology is to improve the sample prediction of multivariate extreme observations. Additionally, a methodology for testing whether this biplot refinement for extremes improves over the traditional biplot is proposed by way of a simulation study.

4.1 THE REFINED PCA BILOT

As stated in Chapter 1, the primary objective of this study is to refine the Traditional PCA biplot methodology to be better suited for multivariate extreme observations. This will be achieved by incorporating the approach described in Section 3.5, by determining an extremal correlation matrix from extreme observations and applying it to the PCA biplot construction methodology. The refinement is obtained for a data matrix X with n observations and p variables as follows:

1. Standardise the data matrix X .

4.1. THE REFINED PCA BILOT

2. Select $k < n$ as the number of tail samples that is used to identify extreme observations.
3. Store the identified extreme observations as \mathbf{X}^* .
4. Use procedure discussed in Section 3.3 to construct a correlation matrix for k tail samples, termed the extremal correlation matrix and denoted as $\mathcal{R}^{extreme}$.
5. Determine the Singular Value Decomposition of $\mathcal{R}^{extreme}$ as,

$$\mathcal{R}^{extreme} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}' \quad (4.1)$$

with,

- $\mathbf{\Lambda}$ a diagonal matrix containing the eigenvalues of $\mathcal{R}^{extreme}$, sorted in descending order.
 - \mathbf{V} a matrix with columns the corresponding eigenvectors, ordered corresponding to its eigenvalues.
6. Using the matrix of eigenvalues $\mathbf{\Lambda}$, find the corresponding diagonal matrix of singular values \mathbf{D} as¹ as,

$$\mathbf{D} = [(\mathbf{\Lambda}^{-1})^{\frac{1}{2}}]$$
 7. Then apply the PCA biplot construction methodology described in Section 2.2 using the derived singular value matrix \mathbf{D} and eigenvector matrix \mathbf{V} obtained in the preceding steps. This will yield the Refined PCA biplot for extreme observations.
 8. Finally, use a different marker or colour to denote extreme observations on the biplot corresponding to observations in \mathbf{X}^*

The biplot obtained through the above methodology is termed the Refined biplot and a biplot obtained through the generic biplot construction methodology is termed the Traditional biplot. The Refined biplot is implemented by modifying the `PCAbipl`

¹Refer to Appendix A.2 for further details.

4.2. SIMULATION DESIGN

function in the R package `UBbipl` developed by le Roux and Lubbe (2013). The functions used to obtain the extremal correlation matrix was derived from code published by Haug *et al.* (2015). The R code to construct the Refined PCA biplot is exhibited in Appendix D. Given that the Refined PCA biplot is constructed based on an extreme correlation matrix, it can be argued that extreme observations are better preserved when performing dimension reduction through PCA and as such extreme observations should better represented by the Refined biplot. However, this would also imply that non-extreme observations will be represented worse. Therefore, a trade-off exists between overall biplot fit and fit for extreme observations. Hence, to investigate if the Refined biplot consistently improves the sample predictivity for extreme observations a simulation study is carried out. The set-up of such a simulation study is described in the subsequent section.

4.2 SIMULATION DESIGN

In this section, a test as to whether the Refined biplot methodology consistently yields better sample predictions for extreme observations compared to the Traditional biplot is presented. This test essentially compares the average sample prediction error for extremes, determined through several simulations of multivariate data, for both approaches. An outline of such a simulation procedure is illustrated by a flowchart in Figure 4.1 on page 47. The flowchart is divided into three subsections or blocks (A), (B), and (C). Each of these subsections will be explained separately as part of the simulation design. Subsection (A) corresponds to the data generating process. Then in subsection (B) the Traditional and Refined biplots are constructed using the generated data. Furthermore, after each of the biplots is constructed the sample prediction error corresponding to extreme observations is measured for both biplots and stored. Finally, in subsection (C) it is tested whether the extreme sample prediction error is significantly lower on average for the Refined biplot than in the case of the Traditional biplot.

4.2. SIMULATION DESIGN

4.2.1 Data generation

The first part of the simulation study denoted (A) in Figure 4.1 is the process of generating data. The simulation study is performed over various independent multivariate datasets. The multivariate distributions the data is generated from use a copula to combine known marginal distributions. Specifically, the Gaussian and Gumbel copulas are applied to couple the marginals. The use of the Gaussian copula is motivated by the fact that the Gaussian copula is known to be tail independent. This would imply that under the Gaussian copula the correlation matrix derived from the observations in the tail is similar to that of the majority of the data. This further implies that the Refined approach would perform poorly since there is no difference in the overall and extreme correlation. Therefore, the data generated using a Gaussian copula will serve as the benchmark case. However, when the Gumbel copula is used to construct a multivariate distribution it has the feature of exhibiting upper tail dependence. The upper tail dependence exhibited by the Gumbel copula would imply that a correlation structure derived from the tails of the data should vary from the correlation of the entire data set. For this case, it is expected that the Refined approach is superior in terms of extreme sample predictivity as compared to the Traditional approach. Algorithms 3.1 and 3.3, as given by Embrechts *et al.* (2002), provide the steps to generate multivariate samples using the Gaussian and Gumbel copulas, respectively.

Using algorithms 3.1 and 3.3, the data generating process is considered for three distinct cases:

- i Multivariate distribution with identical $\text{Gamma}(2, 2)$ marginals and dependence characterised by a Gaussian copula with correlation $\rho_{ij} = 0.5$ for all $i \neq j$.
- ii Multivariate distribution with identical $\text{Gamma}(2, 2)$ marginals and dependence characterised by a Gumbel copula with parameter $\theta = 1.5$.
- iii Multivariate distribution with heterogeneous marginal distributions and dependence characterised by a Gumbel copula with parameter $\theta = 1.5$.

Case (i) is taken as the benchmark case since the Gaussian copula is tail independent.

4.2. SIMULATION DESIGN

This implies that the correlation structure in the tails is similar to that of the entire distribution and such Refined methodology is expected to perform poorly. Case (ii) is considered to examine whether the existence of tail dependence, due to the use of a Gumbel copula, yields better results than Case (i). Furthermore, in Case (iii) different marginals are used with various characteristics coupled with a Gumbel copula. In this case, the performance of the Refined methodology is tested for multivariate data with non-homogeneous marginals.

4.2.2 Traditional and Refined PCA biplot quality measures

In all three of the above cases, the generated multivariate data scaled by its standard deviation and mean, and is used to construct a Traditional biplot and a Refined biplot. Each of the two biplots is constructed for 4-, 5- and 7-variables for cases (i) and (ii), and in case (iii) only 5 variables are considered, with $n = 500$ & $n = 5000$ samples using $k = 20$ & $k = 80$ tail samples. This yields 12 separate datasets to consider and to compare the Traditional and Refined biplots. Hence, a measure is required to test the quality of extreme observations on each of these datasets under the Refined and Traditional biplots. Since only extreme observations are considered the measure of extreme sample prediction error is used to assess the quality of extreme observations under each biplot. Suppose that \mathbf{X}^* is a d -variate matrix of h identified extreme observations. These observations are identified as the observations that hold true for the identity function in (3.39). Further, let \mathbf{x}_i^* be the i^{th} original extreme sample and $\hat{\mathbf{x}}_i^*$ is the i^{th} extreme sample as predicted by the biplot then,

$$\Psi_i^* = [\hat{\mathbf{x}}_i^* - \mathbf{x}_i^*]^2 \quad (4.2)$$

is the extreme sample prediction error for the i^{th} extreme observation. The overall extreme sample error used to compare the respective biplots is determined as the sum over all extreme observations as,

$$\Psi^* = \sum_{i=1}^h \sum_{j=1}^d \Psi_{i,j}^* \quad (4.3)$$

This is denoted as Ψ_T^* and Ψ_R^* for the Traditional and Refined biplot, respectively.

4.3. SUMMARY

4.2.3 Testing procedure

The procedure in (A) and (B) of Figure 4.1 are repeated 100 times, which yield 100 extreme sample prediction errors for each biplot. Next, the average sample prediction error for the Traditional and Refined biplot is determined and denoted as $\bar{\Psi}_T^*$ and $\bar{\Psi}_R^*$, respectively. However, it is not acceptable to just compare whether $\bar{\Psi}_T^* > \bar{\Psi}_R^*$ to conclude that the Refined biplot has a lower average extreme sample error and hence is a better display for extreme observations. In order to make such a conclusion, simulation variation must be taken into account. Hence, the standard error of the simulated extreme sample prediction error is computed and denoted as $S.E.(\Psi_T^*)$ and $S.E.(\Psi_R^*)$ for the Traditional and Refined biplot, respectively. Then to determine whether the average extreme sample prediction error is lower for the Refined biplot than for the Traditional biplot at a 10% significance level, consider the following hypothesis test:

$$H_0 : E[\psi_R^*] - E[\psi_T^*] \geq 0 \quad vs \quad H_1 : E[\psi_R^*] - E[\psi_T^*] < 0 \quad (4.4)$$

The above hypothesis is tested against a significance of $\alpha = 10\%$. The Student-T distribution is used as a test statistic with the p-value calculated as

$$p - value = P \left(T < \frac{\bar{\psi}_R^* - \bar{\psi}_T^*}{\sqrt{S.E.(\psi_R^*)^2 + S.E.(\psi_T^*)^2}} \right) \quad with \quad T \sim t(n-1)$$

The null hypothesis H_0 is rejected if it is found that $p - value < 0.1$. Then it can be said that there is sufficient evidence to conclude that the average extreme sample prediction error under the Refined approach is lower than in the Traditional approach. This means that the Refined biplot methodology provides a superior sample prediction than the Traditional biplot for extreme observations.

4.3 SUMMARY

In this chapter a methodology to construct a biplot that yields superior extreme sample predictions over the traditional biplot methodology was proposed. This new biplot is termed the Refined biplot and is constructed by assuming extreme observations are determined by an elliptical distribution. Furthermore, a procedure was then discussed

4.3. SUMMARY

to test the prediction accuracy of the Refined biplot for extreme observations through a simulation study. The results of this simulation are presented in the subsequent chapter.

4.3. SUMMARY

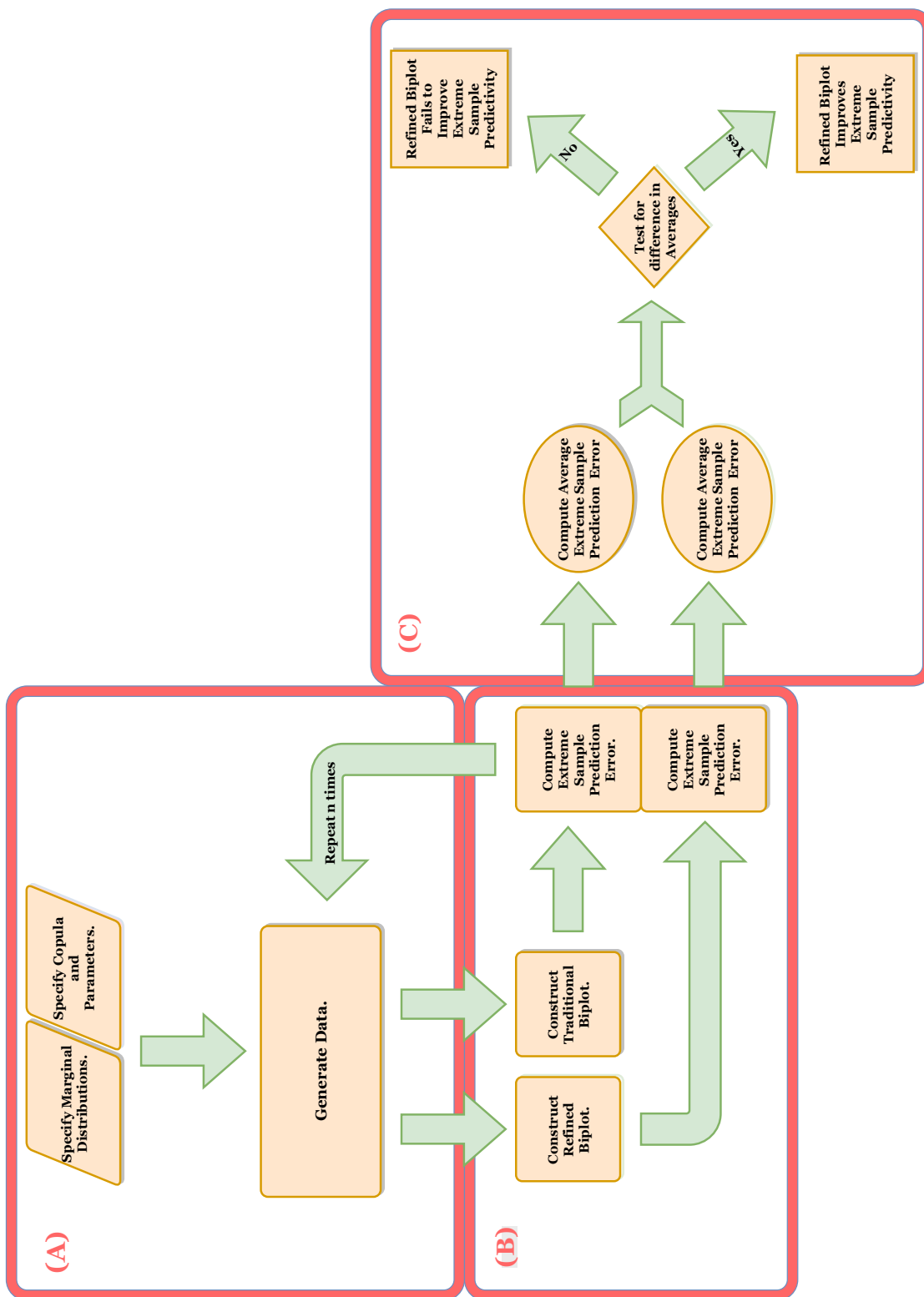


Figure 4.1

Flow diagram of the simulation structure

CHAPTER 5

SIMULATION RESULTS

As discussed in the preceding chapter, the capability of the Refined biplot to improve extreme sample predictivity over the Traditional biplot is tested by way of a simulation study. This chapter is devoted to presenting and explaining the results from the simulation study. The simulation of both biplots' performances is repeated over the following three distinct multivariate distributions:

- i Observations from a multivariate distribution with $Gamma(2, 2)$ underlying marginals which is combined using a Gaussian copula with correlation parameter $\rho_{i,j} = 0.5$ for all $i \neq j$.
- ii Observations from a multivariate distribution with $Gamma(2, 2)$ underlying marginals which is combined using a Gumbel copula with parameter $\theta = 1.5$.
- iii Observations from a multivariate distribution consisting of various marginals which is combined using a Gumbel copula with parameter $\theta = 1.5$.

Each section in this chapter consists of illustrations of both the Traditional and Refined biplot, with accompanied quality measures, constructed from observations simulated from each above-mentioned distributions for the 5-variate case. Additionally, as mentioned in section 4.2.2, all generated data is scaled for each variable by its mean and standard deviation. Further, simulation results for the average extreme sample prediction error for both biplots are given for each of the above distributions under various

5.1. GAUSSIAN COPULA WITH GAMMA MARGINALS

dimensions using a varying number of observations and tail samples.

5.1 GAUSSIAN COPULA WITH GAMMA MARGINALS

In this section, the application of the Refined biplot is presented for a distribution termed from here on as the Gamma-Gaussian distribution. The Gamma-Gaussian distribution is constructed from homogeneous $\text{Gamma}(2,2)$ distributed marginals which are coupled using a Gaussian copula with correlation parameter $\rho = 0.5$ for all variables. In Figure 5.1 a pairs plot consisting histograms for each variable on the diagonal and scatter plots for each pair of variables on the off-diagonal is presented for 5000 observations generated from a 5-variate Gamma-Gaussian distribution. It is clear from the scatter plots that all variables have a positive relationship since the correlation parameter is 0.5 for all variables. Further, from the histograms, all variables are positively skewed. Additionally, as mentioned in section 3.2.1, the Gaussian copula is tail independent which can be observed in the scatter plots since observations in the upper right quadrant of each scatter plot is more sparse indicating a weaker relationship between extreme observations.

5.1. GAUSSIAN COPULA WITH GAMMA MARGINALS

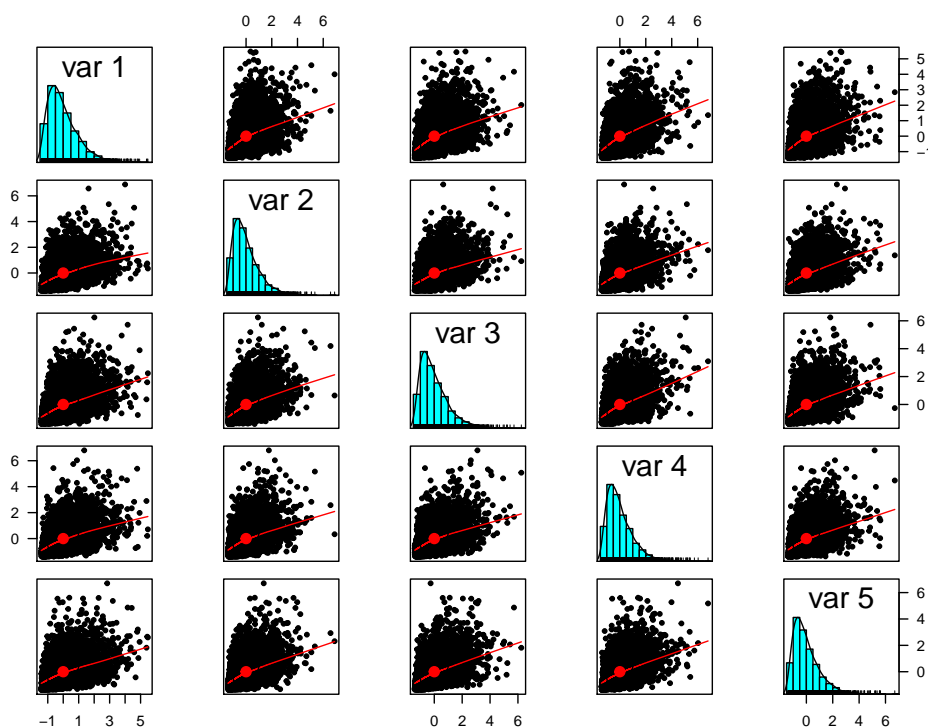
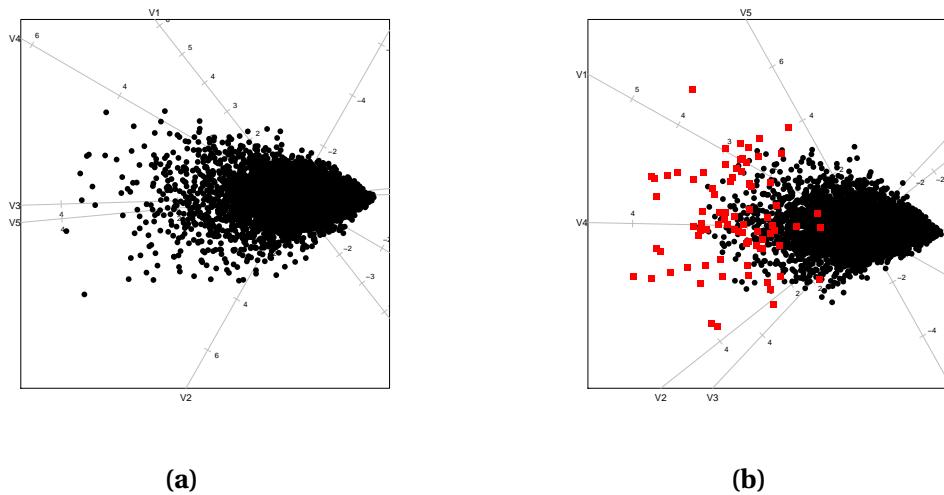


Figure 5.1

Pairs scatter plot of a 5-variate Gamma-Gaussian distribution

Now using the 5000 observations generated from a 5-variate Gamma-Gaussian distribution a Traditional biplot is constructed and is illustrated in Figure 5.2a. Then by arbitrarily choosing the number of tail samples as $k = 80$ the Refined biplot is constructed and is illustrated in Figure 5.2b. The observations depicted as red blocks are the extreme observations used to obtain an extremal correlation matrix which is then used to construct the Refined biplot.

5.1. GAUSSIAN COPULA WITH GAMMA MARGINALS

**Figure 5.2**

Traditional (a) and Refined (b) biplots for 5000 observations from a 5-variate Gamma-Gaussian distribution.

Further, the overall PCA quality for the Traditional biplot is 68.66% and for the Refined biplot is 73.91%. However, it is not enough to compare the PCA quality and further quality measures such as the biplot Predictivity and Adequacy must be considered and are presented in Table 5.1 for each underlying variable. Recall that Adequacy measure assesses how well the biplot axes in 2-dimensional space represent the axes in 5-dimensional space whereas, the measure of predictivity reveals how well each axis predicts the true observations. From Table 5.1 it is clear that the Traditional biplot estimates variables 1, 2 and 4 well owing to it having a higher adequacy and predictivity and for the same reason the Refined biplot estimates variables 3 and 5 better.

5.1. GAUSSIAN COPULA WITH GAMMA MARGINALS

Table 5.1

Predictivity and Adequacy measures for biplots constructed from 5000 observations from a 5-variate Gamma-Gaussian distribution.

Predictivity	1	2	3	4	5
Traditional	0.749	0.898	0.585	0.627	0.581
Refined	0.602	0.695	0.898	0.625	0.876
Adequacy	1	2	3	4	5
Traditional	0.521	0.805	0.203	0.270	0.202
Refined	0.244	0.328	0.507	0.206	0.715

The above measures are important, but for the purpose of this study it is meaningless since only the fit of extreme observations are of interest. Thus, to determine this, the sample prediction error is measured for each variable over all observations and its averages are given in Table 5.2. Subsequently, the observations used to construct the Refined biplot are taken as the extremes for which the extreme sample prediction error is presented for each variable in Table 5.3. Now examining the total overall sample error in Table 5.2 it is clear that the Traditional biplot is superior in sample prediction for all observations. However, when considering only the total extreme sample error in Table 5.3 it appears that the Refined biplot better estimates extreme observations.

Table 5.2

Overall sample error for biplots constructed from 5000 observations from a 5-variate Gamma-Gaussian distribution.

Variable	1	2	3	4	5	Total
Traditional	0.250	0.102	0.415	0.380	0.419	1.567
Refined	0.400	0.359	0.257	0.415	0.151	1.581

5.1. GAUSSIAN COPULA WITH GAMMA MARGINALS

Table 5.3

Extreme sample error for biplots constructed from 5000 observations from a 5-variate Gamma-Gaussian distribution.

Variable	1	2	3	4	5	Total
Traditional	1.687	0.626	2.430	2.438	2.619	9.799
Refined	2.633	2.147	1.465	2.559	0.915	9.721

However the above results are only obtained for one generated dataset and only considers the 5-variate case and in order to ensure that the Refined biplot does improve extreme sample predictions a simulation study must be pursued. The simulation results are presented in Table 5.4 and presents the average extreme and overall sample error for various variable dimensions and under a different choice of n observations and tail samples k . Accompanying, the average sample errors is the standard error of the simulation with p-values for each of the 4, 5, and 7 dimensional cases. It is clear by looking at the p-values in Table 5.4 that in no case does the Refined biplot improve extreme sample predictions over the Traditional biplot. This implies that the extreme sample prediction error given in Table 5.3 is an anomaly and on average the Refined biplot yields worse overall and extreme sample predictions than the Traditional biplot. The reason that the Refined biplot yields worse extreme sample prediction can be attributed to fact that the Gaussian copula is tail independent, meaning there is no difference in the correlation between the extreme observations and that of all the observations. This in part implies that the extremal correlation matrix obtained in the Refined biplot methodology was a substandard estimation of the correlation matrix for all observations.

5.1. GAUSSIAN COPULA WITH GAMMA MARGINALS

Table 5.4

Overall and extreme sample error simulation results for biplots from a Gamma-Gaussian distribution. Each table consists of the average sample error, standard error and p-values for 4, 5 and 7 variables using 500 and 5000 observations for 20 and 80 tail samples.

	4 Variables		5 Variables		7 Variables	
n=5000 k=20	Overall	Extreme	Overall	Extreme	Overall	Extreme
Traditional	1.954	4.138	2.717	5.617	4.175	7.861
(Std.Err)	0.011	0.056	0.009	0.055	0.015	0.092
Refined	1.985	4.252	2.762	5.690	4.235	8.085
(Std.Err)	0.011	0.077	0.010	0.045	0.016	0.089
p-value	0.920	0.804	0.989	0.766	0.972	0.891
n=5000 k=80						
Traditional	1.951	2.669	2.711	3.631	4.171	5.345
(Std.Err)	0.014	0.026	0.008	0.016	0.016	0.037
Refined	1.956	2.675	2.742	3.693	4.197	5.388
(Std.Err)	0.013	0.027	0.009	0.021	0.019	0.041
p-value	0.578	0.543	0.962	0.952	0.771	0.710
n=500 k=20						
Traditional	1.945	4.180	2.714	5.598	4.213	8.193
(Std.Err)	0.015	0.081	0.009	0.060	0.021	0.106
Refined	1.991	4.270	2.765	5.760	4.276	8.408
(Std.Err)	0.013	0.075	0.009	0.063	0.024	0.158
p-value	0.950	0.718	0.997	0.905	0.916	0.790
n=500 k=80						
Traditional	1.943	2.648	2.713	3.612	4.233	5.467
(Std.Err)	0.015	0.028	0.008	0.018	0.023	0.075
Refined	1.962	2.660	2.744	3.671	4.253	5.546
(Std.Err)	0.011	0.025	0.008	0.019	0.023	0.091
p-value	0.769	0.586	0.973	0.946	0.667	0.683

5.2. GUMBEL COPULA WITH GAMMA MARGINALS

5.2 GUMBEL COPULA WITH GAMMA MARGINALS

In this section, the application of the Refined biplot is presented for a distribution termed from here on out as the Gamma-Gumbel distribution. The Gamma-Gumbel distribution is constructed from homogeneous $Gamma(2, 2)$ distributed marginals which are coupled using a Gumbel copula with parameter $\theta = 1.5$. This means that, as mentioned in section 3.2.2., the upper tail dependence coefficient of the distribution is $\lambda_{upper} = 2 - 2^{\frac{1}{1.5}} = 0.851$. This implies that the dependence in the tail is stronger and different to that of the remaining observations not in the tail. A pairs plots are illustrated in Figure 5.3 consisting of a histogram of each underlying variable in the diagonal and in the off-diagonal scatter plots of each pairs of variables is given. This pairs plot is constructed from 5000 observations generated from a 5-variate Gamma-Gumbel distribution. Once again, all variables have a positive relationship. However, in contrast to the Gamma-Gaussian distribution presented in Figure 5.1, extreme observations in the upper right quadrant of each scatter plot are more correlated due to the non-zero upper tail dependence coefficient.

5.2. GUMBEL COPULA WITH GAMMA MARGINALS

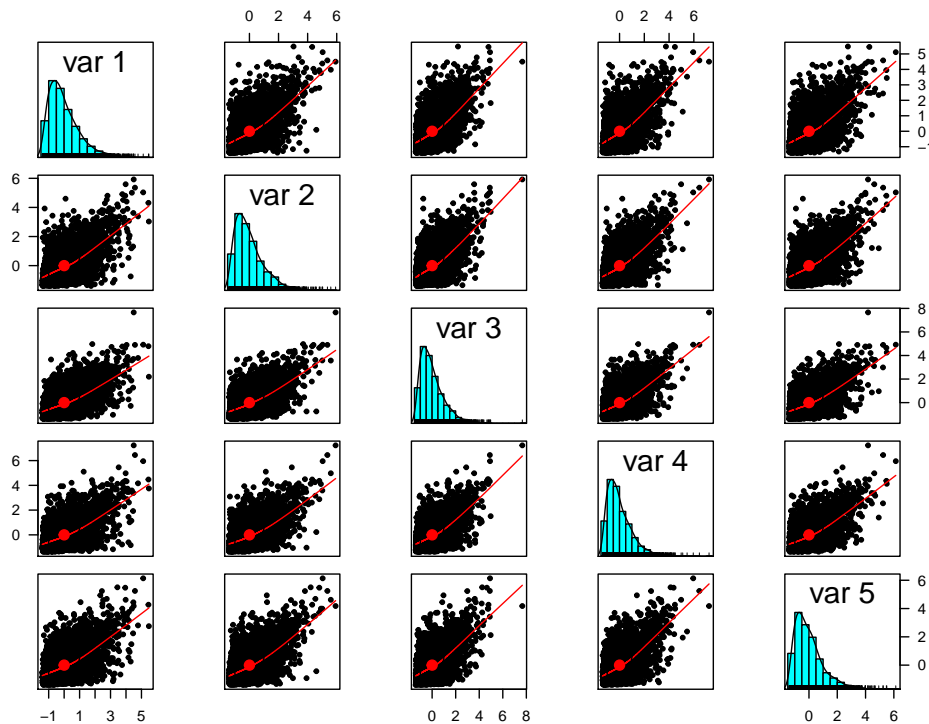
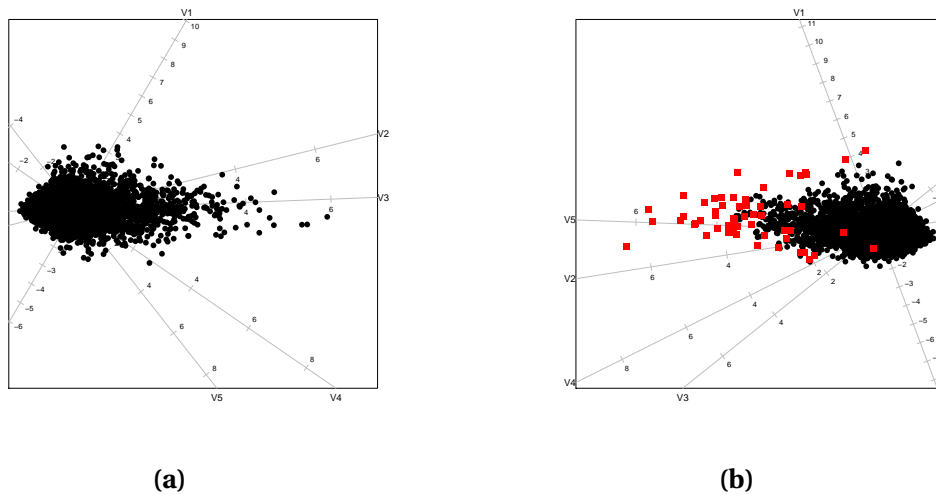


Figure 5.3

Pairs scatter plot of a 5-variate Gamma-Gumbel distribution

Now using the 5000 generated observations from the above 5-variate Gamma-Gumbel distribution a Traditional biplot is constructed and illustrated in Figure 5.4a. Further, by choosing $k = 80$ as the number of tail samples a Refined biplot is constructed and illustrated in Figure 5.4b. Similarly, observations depicted as red blocks are extreme observations used to determine an extremal correlation matrix. Now comparing the biplots in Figure 5.4 with those presented in Figure 5.2, both having the same identical marginals and only differing in copulas, it is clear that the extreme observations in Figure 5.4 are subject to greater dependency due the non-zero tail dependence. This means that extreme observations may yield a vastly different correlation matrix than that of a correlation matrix on all observations.

5.2. GUMBEL COPULA WITH GAMMA MARGINALS

**Figure 5.4**

Traditional (a) and Refined (b) biplots for 5000 observations from a 5-variate Gamma-Gumbel distribution.

Furthermore, the PCA quality for the Traditional and Refined biplots are 75.25% and 77.22%, respectively. Moreover, the Adequacy and Predictivity measures is reported in Table 5.5, which shows that the Refined biplot better represents variables 1 to 4, whereas the Traditional biplot is superior in representing variable 5.

Table 5.5

Predictivity and Adequacy measures for biplots constructed from 5000 observations from a 5-variate Gamma-Gumbel distribution.

Predictivity	1	2	3	4	5
Traditional	0.904	0.681	0.670	0.713	0.808
Refined	0.962	0.751	0.714	0.745	0.688
Adequacy	1	2	3	4	5
Traditional	0.766	0.215	0.201	0.296	0.522
Refined	0.922	0.246	0.327	0.283	0.222

5.2. GUMBEL COPULA WITH GAMMA MARGINALS

Consider now the average overall and extreme sample prediction error in Table 5.6 and 5.7. It is noted that, although the Refined biplot yield a poorer overall sample prediction, it does provide improved extreme sample prediction given that its total average extreme sample error is 3.427 which is lower than 3.598 for the Traditional biplot. This provides some evidences that the Refined biplot may better present extreme samples than the Traditional biplot. However simulation is required to test whether this assertion holds consistently over various dimensions.

Table 5.6

Overall sample error for biplots constructed from 5000 observations from a 5-variate Gamma-Gumbel distribution.

Variable	1	2	3	4	5	Total
Traditional	0.096	0.318	0.330	0.287	0.192	1.224
Refined	0.033	0.309	0.283	0.295	0.323	1.244

Table 5.7

Extreme sample error for biplots constructed from 5000 observations from a 5-variate Gamma-Gumbel distribution.

Variable	1	2	3	4	5	Total
Traditional	0.319	0.947	1.000	0.775	0.558	3.598
Refined	0.094	0.926	0.811	0.820	0.777	3.427

The results of the simulated average overall and extreme sample prediction error are presented with standard errors and p-values in Table 5.8. Considering that the p-values are tested at a 10% significance level. It is immediately clear that the Refined biplot improves extreme sample prediction in the 4- and 5-variate case for $n = 5000$ observations with $k = 80$ tail samples. However, this does not hold in the 7-variate case. Conversely, if the choice of $k = 80$ is used with $n = 5000$ observations then it can not be said at a 10% significance level that the Refined biplot is superior to the Traditional

5.2. GUMBEL COPULA WITH GAMMA MARGINALS

biplot in representing extreme samples. In the final two tables of Table 5.8, $n = 500$ observations is used in the simulation of biplot extreme sample prediction error. When $k = 20$ tail samples is considered then the Refined biplot only improves the extreme sample prediction over the Traditional biplot in the 4-variate case. Finally, the Refined biplot fails to improve extreme sample prediction error in all dimensions when $k = 80$ tail samples is considered.

5.2. GUMBEL COPULA WITH GAMMA MARGINALS

Table 5.8

Overall and extreme sample error simulation results for biplots from a Gamma-Gumbel distribution. Each table consists of the average sample error, standard error and p-values for 4, 5 and 7 variables using 500 and 5000 observations for 20 and 80 tail samples.

	4 Variables		5 Variables		7 Variables	
n=5000 k=20	Overall	Extreme	Overall	Extreme	Overall	Extreme
Traditional	1.732	2.600	2.458	3.406	3.828	5.010
(Std.Err)	0.020	0.056	0.009	0.043	0.026	0.040
Refined	1.716	2.401	2.457	3.293	3.894	5.196
(Std.Err)	0.016	0.062	0.008	0.044	0.026	0.035
p-value	0.323	0.048	0.476	0.097	0.893	0.993
n=5000 k=80						
Traditional	1.774	1.969	2.473	2.667	3.866	3.981
(Std.Err)	0.013	0.030	0.014	0.025	0.028	0.058
Refined	1.777	1.912	2.472	2.616	3.891	4.049
(Std.Err)	0.026	0.026	0.012	0.028	0.034	0.062
p-value	0.535	0.159	0.495	0.170	0.654	0.715
n=500 k=20						
Traditional	1.797	2.539	2.452	3.435	3.814	4.959
(Std.Err)	0.024	0.065	0.013	0.067	0.038	0.218
Refined	1.813	2.362	2.503	3.403	3.906	5.178
(Std.Err)	0.016	0.051	0.012	0.069	0.039	0.211
p-value	0.653	0.065	0.977	0.409	0.883	0.695
n=500 k=80						
Traditional	1.814	2.021	2.465	2.649	3.835	3.927
(Std.Err)	0.026	0.053	0.014	0.028	0.032	0.068
Refined	1.831	2.037	2.495	2.618	3.869	3.973
(Std.Err)	0.022	0.052	0.014	0.027	0.026	0.074
p-value	0.642	0.558	0.853	0.286	0.719	0.628

5.3. GUMBEL COPULA WITH HETEROGENEOUS MARGINALS

From the simulation results presented in Table 5.8 it can be concluded that the Refined biplot methodology improves the extreme sample predictivity, only if the underlying multivariate distribution has non-zero tail dependence. Further, if a large number of observations is combined with a small number of tail samples, as considered in the case with $n = 5000$ and $k = 20$, then the Refined methodology yields superior extreme sample predictions for up to 5 dimensional datasets. However, the Refined biplot methodology is inferior to the Traditional biplot methodology when a smaller set of observations is used with the number of tail samples constituting a large fraction of the total observations, as in the case with $n = 500$ and $k = 80$.

5.3 GUMBEL COPULA WITH HETEROGENEOUS MARGINALS

In this final section the application of the Refined biplot is presented for a multivariate distribution termed the Heterogeneous-Gumbel distribution. The Heterogeneous-Gumbel distribution is a 4-variate distribution consisting of arbitrary marginals distributions $Gamma(2,2)$, $Beta(1,2)$, $Student - t(5)$, and $Gumbel(0,2)$. The marginals are combined using a Gumbel copula with parameter $\theta = 1.5$. As in the preceding section, this implies that the Heterogeneous-Gumbel distribution has a non-zero upper tail dependence coefficient of $\lambda_{upper} = 0.851$. The goal in this section is to determine whether the Refined biplot yield better extreme sample prediction even if underlying variables have different marginals, while still being subject to upper tail dependence. Once again a pairs plot is illustrated in Figure 5.5, which is constructed from 5000 observations generated from the 4-variate Heterogeneous-Gumbel distribution. Further, the histograms indicates that each variable originates from a different marginal distribution. Additionally, since the distribution has non-zero upper tail dependence, the upper right quadrant of each scatter plot exhibits a stronger dependence.

5.3. GUMBEL COPULA WITH HETEROGENEOUS MARGINALS

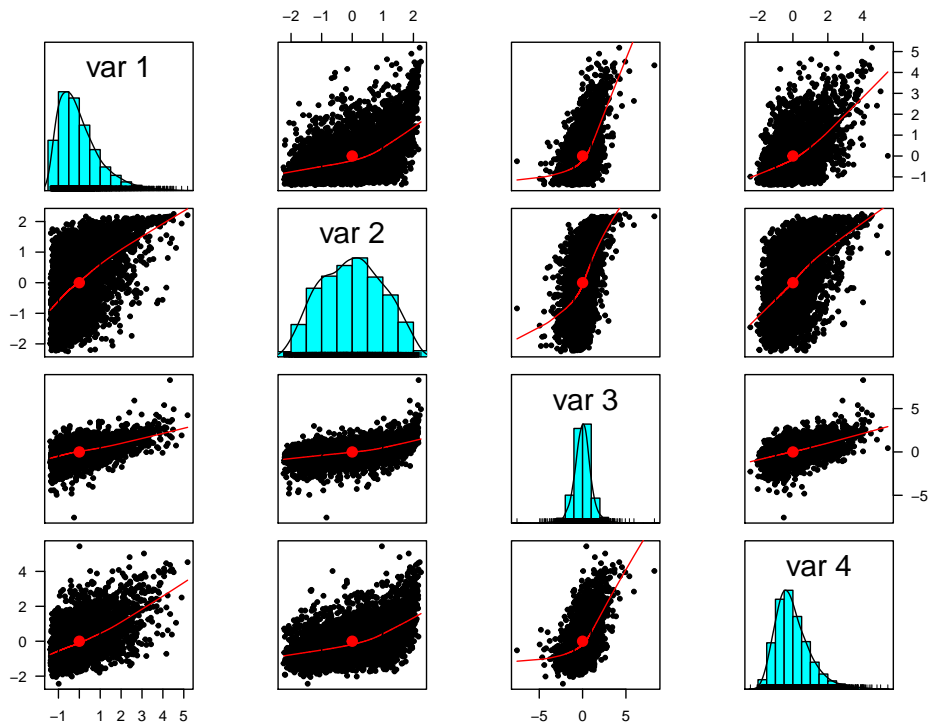


Figure 5.5

Pairs scatter plot of a 4-variate Heterogeneous-Gumbel distribution

Using the generated 5000 observations from the 4-variate Heterogeneous-Gumbel distribution a Traditional biplot is constructed and presented in Figure 5.6a. Furthermore, by selecting the tail sample to be $k = 80$, a Refined biplot can be constructed and is presented in Figure 5.6b. Similarly, observations depicted by red blocks are the extreme observations used to construct the Refined biplot. Due to the observations being generated by a distribution constructed from a variety of marginals the shape of the data on the biplot is somewhat skewed. Lastly, the orientations of the axes are different for the two biplots, showing that different correlation matrices are used to construct the respective biplots.

5.3. GUMBEL COPULA WITH HETEROGENEOUS MARGINALS

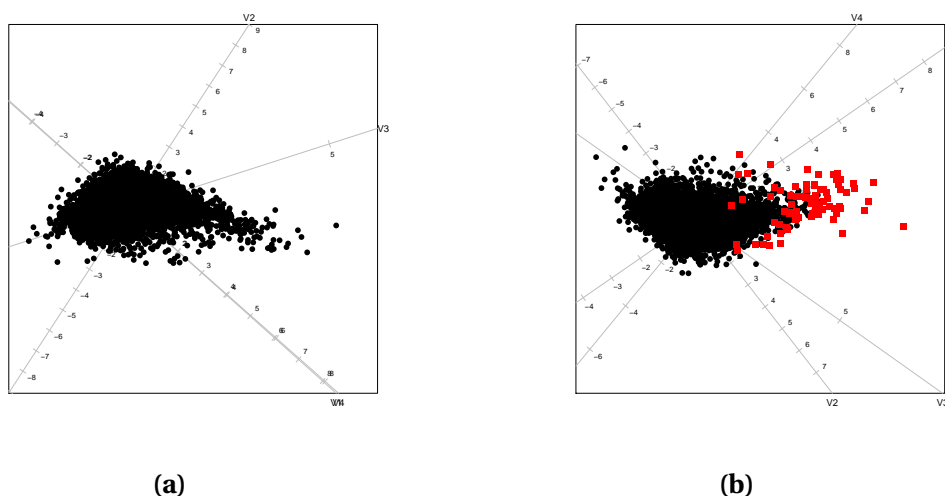


Figure 5.6

Traditional (a) and Refined (b) biplots for 5000 observations from a 4-variate Heterogeneous-Gumbel distribution.

The PCA quality of the Traditional and Refined biplot is 74.87% and 78.67%, respectively. Moreover, the Adequacy and Predictivity measures is reported in Table 5.9 which shows that the Traditional biplot perfectly predicts variable 2 and yield superior predictivity for variable 4. The Refined biplot on the other hand yield superior predictivity for variables 1 and 3.

Table 5.9

Predictivity and Adequacy measures for biplots constructed from 5000 observations from a 4-variate Heterogeneous-Gumbel distribution.

Predictivity	1	2	3	4
Traditional	0.682	1	0.641	0.672
Refined	0.746	0.859	0.896	0.645
Adequacy	1	2	3	4
Traditional	0.338	1	0.328	0.334
Refined	0.31	0.675	0.742	0.273

5.3. GUMBEL COPULA WITH HETEROGENEOUS MARGINALS

Considering the average overall and extreme sample prediction error in Table 5.10 and 5.11. It can be seen that the Traditional biplot performs slightly better in overall predictions. However, the Refined biplot significantly reduces the extreme sample error from 3.542 in the Traditional biplot to 3.117. Hence, there is some evidence showing that the Refined biplot may yield better extreme sample predictions than the Traditional biplot. In order to show that this improvement generally holds the above process in simulated 100 times for 5000 observations generated from the 4-variate Heterogeneous-Gumbel distribution.

Table 5.10

Overall sample error for biplots constructed from 5000 observations from a 4-variate Heterogeneous-Gumbel distribution.

Variable	1	2	3	4	Total
Traditional	0.318	0.000	0.359	0.328	1.005
Refined	0.339	0.177	0.140	0.362	1.017

Table 5.11

Extreme sample error for biplots constructed from 5000 observations from a 4-variate Heterogeneous-Gumbel distribution.

Variable	1	2	3	4	Total
Traditional	1.165	0.000	1.199	1.178	3.542
Refined	1.079	0.452	0.351	1.235	3.117

5.3. GUMBEL COPULA WITH HETEROGENEOUS MARGINALS

Table 5.12

Overall sample error simulation results for biplots from a 4-variate Heterogeneous-Gumbel distribution. The table constitutes the average sample error, standard error and p-values using 5000 observations and 80 tail samples.

Variable	1	2	3	4	Total
Traditional	0.528	0.256	0.486	0.540	1.810
(Std.Err)	0.005	0.016	0.013	0.004	0.008
Refined	0.470	0.483	0.475	0.476	1.904
(Std.Err)	0.012	0.015	0.013	0.012	0.007
p-value	0.000	1.000	0.344	0.000	1.000

Table 5.13

Extreme sample error simulation results for biplots from a 4-variate Heterogeneous-Gumbel distribution. The table constitutes the average sample error, standard error and p-values using 5000 observations and 80 tail samples.

Variable	1	2	3	4	Total
Traditional	0.786	0.174	0.538	0.773	2.272
(Std.Err)	0.021	0.018	0.025	0.018	0.036
Refined	0.623	0.440	0.481	0.581	2.125
(Std.Err)	0.030	0.027	0.021	0.027	0.034
p-value	0.001	1.000	0.108	0.000	0.019

The Table 5.12 the simulated average overall prediction error is reported and Table 5.13 provides the results for the simulated average extreme prediction error. It is clear that the Refined biplot is inferior in the overall prediction since the p-value is found to be 1 in Table 5.12. Conversely, for extreme sample predictions the p-value is determined to

5.4. SUMMARY

be 0.019 meaning that even at a 5% significance level there is sufficient evidence to say that the Refined biplot better predicts extreme samples than the Traditional biplot.

5.4 SUMMARY

The aim of this chapter was to test whether the Refined biplot methodology improves sample predictivity over the Traditional biplot for extreme observations. To test this assertion, a simulation procedure was implemented and results of this procedure were reported in this chapter. This procedure was reported for three constructed multivariate distributions termed the Gamma-Gaussian, Gamma-Gumbel and Heterogeneous-Gumbel distributions. For the Gamma-Gaussian distribution it is concluded that the Refined biplot fails to improve on the Traditional biplot when predicting extreme observations. This lack of improvement was attributed to the tail independence property of the Gaussian copula. This was affirmed when the results for Gamma-Gumbel distribution, which is not tail independent, was studied which showed that the Refined biplot does improve extreme sample predictivity. However, this improvement was only when moderate dimensional data was considered such as 4- or 5-variate distributions. Furthermore, an improvement is yielded only when the ratio of tail samples to number of observations was small. Finally, the Refined biplot was applied to data generated from a multivariate distribution with heterogeneous marginals combined through a Gumbel copula. The reason for considering such a distribution was to test whether the Refined biplot improves extreme sample predictivity when tail dependence is assumed but marginals differ. Under these circumstances it was determined that the Refined methodology was superior to the Traditional methodology in the 4-variate case.

CHAPTER 6

FINANCIAL APPLICATION

The credibility of the Refined biplot was proved on simulated data in the previous chapter. In this chapter, however, the Refined biplot is examined on real-world financial data. The fact that financial returns are known to be fat-tailed would imply that the Refined biplot should improve the prediction of extreme observations of a series of financial returns.

6.1 FOREIGN EXCHANGE APPLICATION

More specifically, the Refined biplot is applied to Rand foreign currency exchange rate monthly returns for the period July 2013 to June 2018. The original observations are given in Table B.1 with the corresponding standardised observations presented in Table B.2. The foreign currencies considered are the Euro, Yen, UK Sterling, US dollar, Australian dollar, Rupee and Swiss Franc. Thus the dataset used contains 7 variables and 60 observations.

Similar to the previous chapter, a pairs plot of returns is presented in Figure 6.1. Since only 60 observations are considered, no conclusions of upper tail dependencies can be seen on the scatter plots, other than the variables appear to have a positive relationship. Furthermore, the histograms on the diagonals show that each variable has a somewhat symmetric distribution.

6.1. FOREIGN EXCHANGE APPLICATION

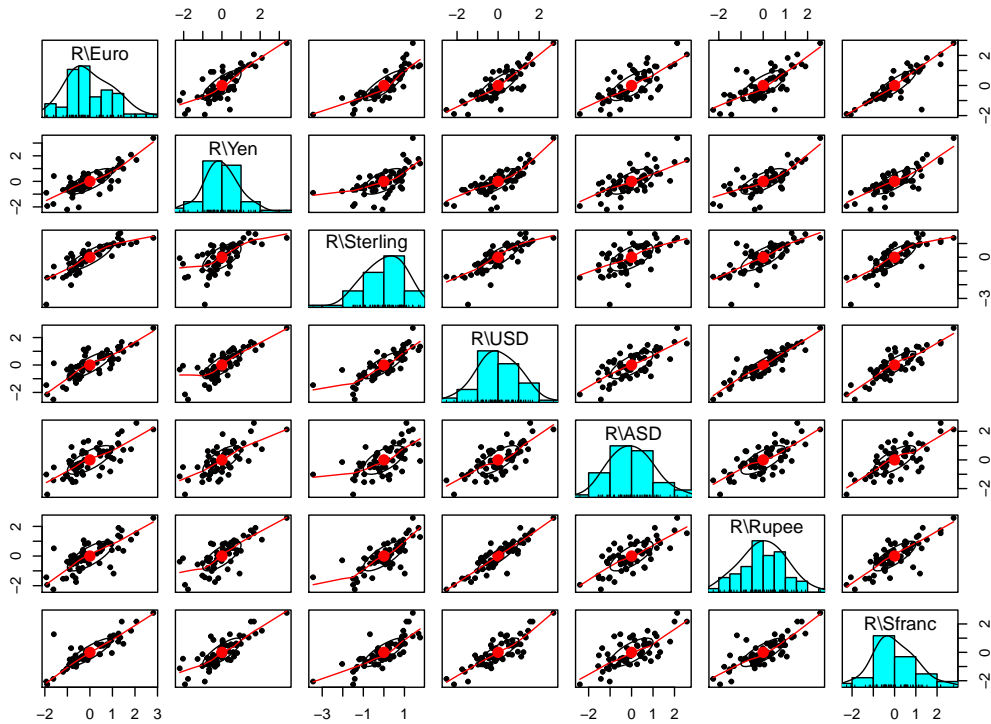


Figure 6.1

Pairs scatter plot for Rand/Foreign currency exchange rate monthly returns

Choosing the number of tail samples to be $k = 10$, then a Refined biplot is constructed and presented in Figure 6.3 with the corresponding Traditional biplot in Figure 6.2. In a similar fashion, the observations denoted by red blocks in the Refined biplot, correspond to the extreme observations used to construct the Refined biplot. Once again the change in the orientations of the axes indicate that different correlation matrices are used to construct each the respective biplots.

The PCA quality is determined to be 94.55% and 86.73% for the Refined and Traditional biplots, respectively. Similar to the previous chapter, the Predictivity and Adequacy are reported in Table 6.1. However the Predictivity and Adequacy are of less importance since the only measure of interest is the extreme sample predictivity.

6.1. FOREIGN EXCHANGE APPLICATION

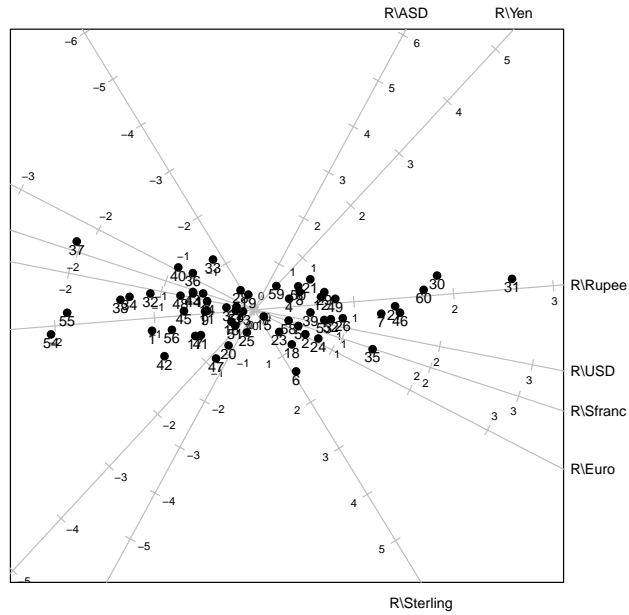


Figure 6.2

Traditional biplot for Rand/Foreign currency exchange rate monthly returns

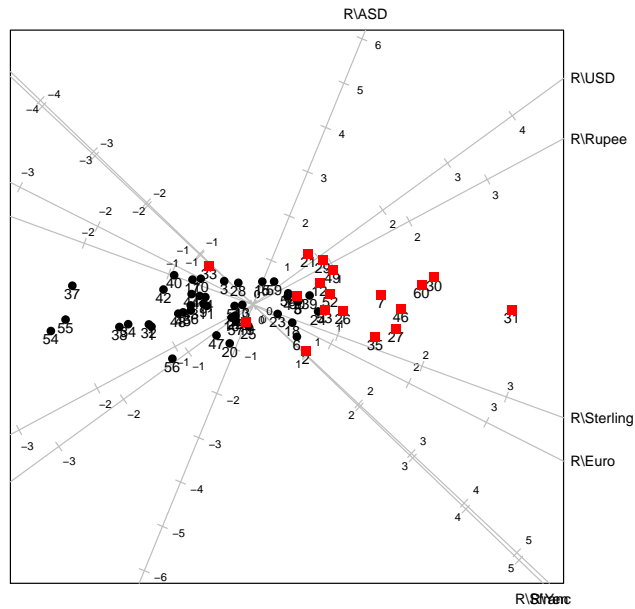


Figure 6.3

Refined biplot for Rand/Foreign currency exchange rate monthly returns

6.1. FOREIGN EXCHANGE APPLICATION

Table 6.1

Predictivity and Adequacy measures for biplots constructed using Rand/Foreign currency exchange rate monthly returns

Predictivity	R\Euro	R\Yen	R\Sterling	R\USD	R\ASD	R\Ruppee	R\Sfranc
Traditional biplot	0.870	0.821	0.925	0.892	0.868	0.821	0.875
Refined biplot	0.949	0.981	0.973	0.945	0.942	0.899	0.931
Adequacy							
Traditional biplot	0.191	0.290	0.500	0.164	0.539	0.146	0.170
Refined biplot	0.196	0.283	0.184	0.230	0.646	0.189	0.273

In Table 6.2 and 6.3 the average overall and extreme sample prediction errors are reported. It is clear to see that for overall predictions the Traditional biplot is superior since it yields a lower sample prediction error of 0.913 compared to 1.024 for the Refined biplot. However, the Refined biplot significantly reduces the extreme sample error from 1.281 to 1.197. This implies that the Refined biplot is more suited to extreme observations than the Traditional biplot.

Table 6.2

Average overall sample prediction error for Rand/Foreign currency exchange rate monthly returns

Variable	R\Euro	R\Yen	R\Sterling	R\USD	R\ASD	R\Ruppee	R\Sfranc	Total
Traditional biplot	0.128	0.176	0.073	0.107	0.130	0.176	0.123	0.913
Refined biplot	0.123	0.222	0.208	0.110	0.105	0.153	0.102	1.024

6.1. FOREIGN EXCHANGE APPLICATION

Table 6.3

Average extreme sample prediction error for Rand/Foreign currency exchange rate
monthly returns

Variable	R\Euro	R\Yen	R\Sterling	R\USD	R\ASD	R\Ruppee	R\Sfranc	Total
Traditional biplot	0.267	0.194	0.082	0.080	0.177	0.240	0.242	1.281
Refined biplot	0.262	0.115	0.220	0.108	0.099	0.209	0.184	1.197

Lastly, the predictive accuracy of a single observation is considered corresponding to the return during January 2019. As discussed in section 2.3, observations can be estimated from the biplot by drawing orthogonal lines between each of the axes and the observation. This is done and illustrated in Figures 6.4 and 6.5, with the original values and those obtained from each biplot reported in Table 6.4. Upon further inspection of Table 6.4, it is clear that the Refined biplot yields better estimates for the exchange rate returns of the *R\Euro*, *R\ASD*, *R\Ruppee*, *R\Franc*, with the Traditional biplot yielding better estimates for the other currencies.

6.1. FOREIGN EXCHANGE APPLICATION

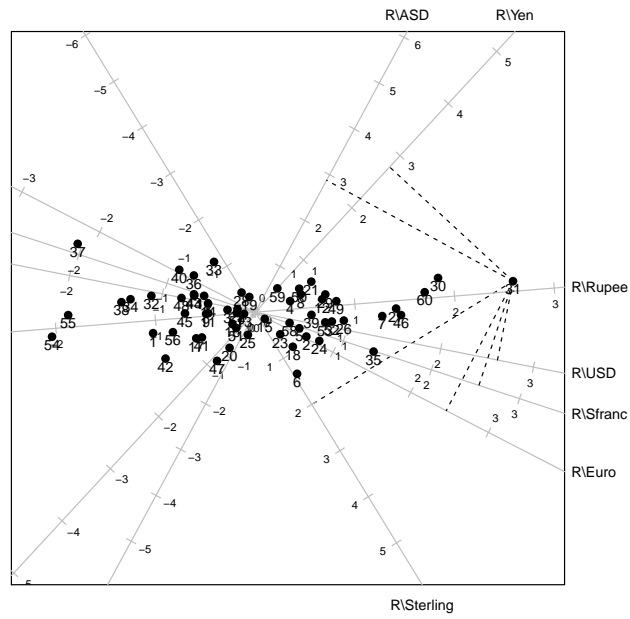


Figure 6.4

Traditional biplot prediction Rand/Foreign currency exchange rate monthly return on
January 2016

6.2. SUMMARY

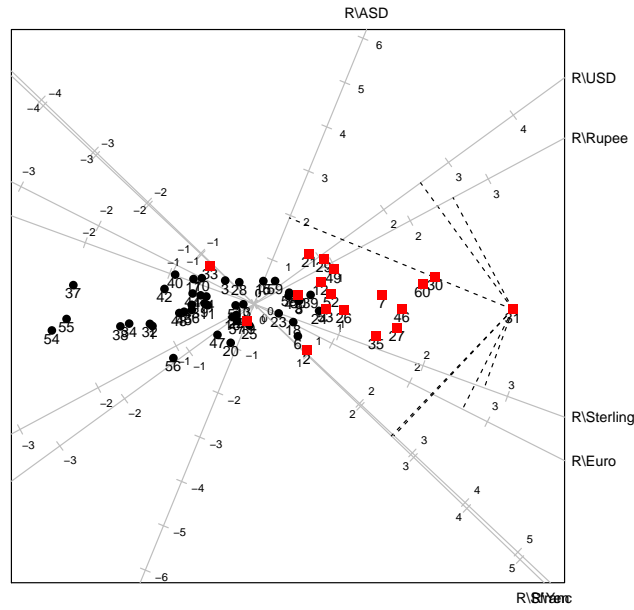


Figure 6.5

Traditional biplot prediction Rand/Foreign currency exchange rate monthly return on January 2016

Table 6.4

Prediction of Rand/Foreign currency exchange rate monthly return on January 2016

Prediction	R\Euro	R\Yen	R\Sterling	R\USD	R\ASD	R\Rupee	R\franc
Original sample	2.818	3.391	1.409	2.698	2.127	2.572	2.780
Traditional biplot estimate	2.457	2.784	2.000	2.611	2.879	2.594	2.539
Refined biplot estimate	2.681	2.652	2.725	2.568	1.946	2.549	2.579

6.2 SUMMARY

The purpose of this chapter was to investigate whether the Refined biplot can be applied to real-world financial data. The data considered is monthly returns of seven Rand foreign currency exchange rates. The Refined biplot demonstrated superior prediction accuracy for extreme observations. The results in this chapter provide some

6.2. SUMMARY

evidence that the Refined biplot can be used together with the Traditional biplot as a tool that provides superior visualisation of financial data that is subject to extremes.

CHAPTER 7

CONCLUSION AND OPEN QUESTIONS

As stated in Chapter 1, one of the objectives of this study was to review the theoretical background of Principal Component Analysis (PCA) biplots and highlight its practical application. This was done in the first part of the literature review, in Chapter 2, where PCA biplots were the focus of the discussion. At the outset of Chapter 2, the development of the biplot methodology to visualise multivariate data was briefly discussed. The mathematical background of PCA was then discussed in detail with its application to construct biplots. Furthermore, an illustration of a biplot was presented and the interpretation of biplots was then explained by way of an example. The chapter was concluded with a discussion of various approaches to evaluate the fit of a biplot.

A proxy method for constructing a biplot is to standardise the underlying dataset and apply PCA to the correlation matrix. This implies that a further investigation of methods to characterise dependence may be beneficial. This was done in Chapter 3, which started by presenting various empirical dependence measures that can be used as alternatives to correlation. The focus of this chapter was mainly on the application of copulas to characterise dependence. In addition to the background regarding copulas, a detailed examination of various copula families and their underlying properties was done. In particular, the background and properties of the Elliptical and Archimedean copula families were discussed. Ultimately, Chapter 3 was concluded by further analysing copula-based methods to analyse dependence for multivariate extremes. A proposed method was to fit an Elliptical copula to extremes and then derive

a correlation matrix characterised by extreme observations. This same methodology was later applied to adjust biplots in order to accommodate multivariate extremes.

The main objective of this study was to propose a methodology to adjust biplots in order to improve prediction accuracy for multivariate extreme observations. To achieve this a correlation matrix for extreme observations was first estimated using the approach presented at the end of Chapter 3. Then by applying PCA to this correlation matrix, a biplot can be constructed that was proposed to be more suited for extreme observations. The biplot derived through this methodology was termed the Refined biplot and the original biplot methodology was termed the Traditional biplot. The second part of Chapter 4 presented a methodology to test whether the Refined biplot improved extreme sample prediction error over the Traditional biplot. This procedure was executed through a simulation of both biplots on the same generated dataset from which the extreme sample prediction error was measured. The aim of the simulation procedure was to test whether the average extreme sample prediction error was significantly lower for the Refined biplot than for the Traditional biplot.

The results from the simulation procedure discussed in Chapter 4 was then reported in Chapter 5. The multivariate data used for the simulation procedure were generated from three constructed multivariate distributions. The first distribution was termed the Gamma-Gaussian distribution and composed of homogeneous $Gamma(2,2)$ marginals combined by way of a Gaussian copula. The results yielded from this distribution showed no improvement for the Refined biplot methodology. This was mainly attributed to the Gaussian copula being tail independent. The second distribution was termed the Gamma-Gumbel distribution which composes of homogeneous $Gamma(2,2)$ marginals combined using a Gumbel copula. The use of the Gumbel copula means that the assumption is made of non-zero upper tail dependence. As a result, the Refined biplot was found to improve extreme sample prediction accuracy over the Traditional biplot. However, improvement was only found for datasets of up to 5 dimensions and only in cases where the proportion of tail samples compared to the number of observations was small. The final distribution was termed the Heterogeneous-Gumbel distribution and was comprised of the following heterogeneous marginals, namely, the

$\text{Gamma}(2,2)$, $\text{Beta}(1,2)$, $\text{Student}-t$ with 5 degrees of freedom, and the $\text{Gumbel}(0,2)$ distributions. These distributions were combined through a Gumbel copula which again leads to non-zero upper tail dependence. Since various marginals were specified, for comparison's sake, only the 4-variate case was considered. That being the case, it was shown that the Refined biplot does significantly improve the extreme sample prediction accuracy.

Finally, the application of the Refined biplot was evaluated on real-world financial data. The data considered were monthly returns on seven Rand/Foreign currency exchange rates. The foreign currencies considered were the Euro, Yen, UK Sterling, US dollar, Australian dollar, Rupee and Swiss franc which were observed over the period July 2013 to June 2018. This yielded a dataset containing 7 variables and 60 observations. It was then shown that the Refined biplot does yield superior extreme sample prediction accuracy than the Traditional biplot. Hence, there is some evidence that the Refined biplot methodology can be applied to financial data that is subject to extremes.

The main contribution of this study is the proposed Refined biplot methodology. The aim of the development of this methodology is to improve extreme sample prediction accuracy in biplots. It is shown in the study that under suitable conditions the Refined biplot does improve extreme sample prediction accuracy over the Traditional biplot. These conditions are that the underlying dataset should have non-zero tail dependence and the proportion of tail samples to total observations should be small. Additionally, since the Refined biplot always decreases the overall sample prediction accuracy, it is proposed to be used as an addition to the Traditional biplot if conditions are suitable. Thus, both the Refined and Traditional biplot should be used together as a visualisation tool.

This study serves only as a proof of concept for the Refined biplot methodology and, as such, opens many further research areas. Regarding this study, further research is required to determine which characteristics, other than tail dependence, a dataset should have in order for the Refined biplot to be useful. Furthermore, an investigation is required for optimally choosing the number of tail samples that yields the best

extreme sample prediction accuracy. Admittedly, the use of a correlation matrix is arbitrary due to the fact that it is a mere product of a covariance matrix estimated from data scaled by its standard deviation. This implies that if some other scaling parameter can be used that is related to the extreme value characteristic of the underlying dataset it may yield better extreme sample prediction accuracy using the Traditional biplot methodology.

APPENDIX A

REQUIRED LINEAR ALGEBRA RESULTS

A.1 Spectral decomposition

Given a symmetric matrix $A: n \times n$, with $rank(A) = r \leq n$, its spectral decomposition is given by:

$$A = \mathbf{V}\mathbf{D}\mathbf{V}' = \sum_{i=1}^n \lambda_i \mathbf{v}_i \mathbf{v}_i' \quad (\text{A.1})$$

where,

- i. $\mathbf{V}: n \times n$ is a orthogonal matrix , with column vectors \mathbf{v}_i the normalised eigenvectors of A .
- ii. $\mathbf{D}: n \times n$ is a diagonal matrix, with non-zero diagonal elements equal to the eigenvalues of A , i.e. $[\mathbf{D}_{ii}] = \lambda_i$.

Since $rank(A) = r$ the last $n - r$ diagonal elements of \mathbf{D} equals zero. Further, it is assumed throughout this study that the diagonal elements in \mathbf{D} is ordered from largest to smallest i.e. $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r$ with corresponding eigenvectors of \mathbf{V} ordered accordingly.

A.2 Singular value decomposition (SVD)

Given a centred data matrix $\mathbf{X} : n \times p$, with $\text{rank}(\mathbf{X}) = r \leq p \leq n$, consisting of n observations and p variables. The singular value decomposition (SVD) of \mathbf{X} is given by:

$$\mathbf{X} = \mathbf{U}\mathbf{\Omega}\mathbf{V}' \quad (\text{A.2})$$

with,

- i. $\mathbf{U} : n \times r$ and $\mathbf{V} : p \times r$ are orthonormal matrices such that $\mathbf{U}'\mathbf{U} = \mathbf{I}_r$ and $\mathbf{V}'\mathbf{V} = \mathbf{I}_r$.
- ii. $\mathbf{\Omega} : r \times r$ is a diagonal matrix.

Further, the SVD of $\mathbf{X}'\mathbf{X} : p \times p$ is given by:

$$\mathbf{X}'\mathbf{X} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}' \quad (\text{A.3})$$

with,

- i. $\mathbf{\Lambda} : r \times r = \mathbf{\Omega}'\mathbf{\Omega} = \mathbf{\Omega}^2$.

The sample covariance matrix is derived as $\mathbf{S}^2 = \frac{1}{n-1}\mathbf{X}'\mathbf{X}$ with its SVD given by:

$$\mathbf{S}^2 = \mathbf{V}\mathbf{\Lambda}^*\mathbf{V}' \quad (\text{A.4})$$

with,

- i. $\mathbf{\Lambda}^* : r \times r = \frac{1}{n-1}\mathbf{\Lambda}$

This implies from (A.2) that $\mathbf{D} = [(n-1)\mathbf{\Lambda}^*]^{\frac{1}{2}}$.

The non-zero entries in the diagonal matrix \mathbf{D} are referred as the singular values of \mathbf{X} , denoted as d_i , $i = 1, \dots, r$. Similarly, columns vectors of the matrix \mathbf{V} are termed the singular vectors of \mathbf{X} . Singular vectors are uniquely defined up to a multiplication of -1 . Finally, spectral decomposition is a special case of SVD such that the spectral decomposition of \mathbf{S}^2 yields the same eigenvector matrix \mathbf{V} and singular values are related to eigenvalues as $\lambda_i = (n-1)d_i^2$.

APPENDIX B

DATA SETS

This appendix lists all the data sets used in this study.

Table B.1

Rand foreign currency exchange rate monthly returns for period July 2013 to June 2018.

Date	R\Euro	R\Yen	R\Sterling	R\USD	R\ASD	R\Ruppee	R\Sfranc
2013/07/31	-2.14%	-3.36%	-3.23%	-1.36%	-4.07%	-3.82%	-2.44%
2013/08/31	3.59%	3.56%	3.87%	1.89%	0.27%	-3.40%	3.84%
2013/09/30	-0.77%	-2.36%	1.26%	-1.00%	1.74%	-2.23%	-0.87%
2013/10/31	1.55%	0.89%	0.90%	-0.69%	1.87%	2.87%	1.85%
2013/11/30	1.73%	0.49%	2.80%	2.81%	0.85%	1.24%	1.66%
2013/12/31	3.13%	-1.78%	3.41%	1.62%	-2.07%	2.83%	3.64%
2014/01/31	4.23%	4.30%	5.26%	4.78%	3.22%	4.45%	3.66%
2014/02/28	1.22%	2.92%	1.52%	1.02%	2.34%	0.72%	2.19%
2014/03/31	-0.94%	-2.44%	-1.75%	-2.20%	-0.98%	-0.12%	-0.73%
2014/04/30	-1.99%	-2.21%	-1.22%	-1.88%	0.78%	-0.84%	-2.05%
2014/05/31	-2.00%	-0.59%	-0.75%	-1.42%	-1.56%	0.37%	-2.13%
2014/06/30	1.67%	2.32%	2.96%	2.64%	3.25%	1.87%	1.89%
2014/07/31	-0.49%	0.19%	0.92%	-0.12%	0.10%	-0.66%	-0.24%
2014/08/31	-1.66%	-1.15%	-2.17%	0.03%	-0.80%	-1.33%	-1.42%

Continued on next page

Table B.1 – continued from previous page

Date	R\Euro	R\Yen	R\Sterling	R\USD	R\ASD	R\Ruppee	R\Sfranc
2014/09/30	-0.42%	-1.36%	0.27%	2.68%	0.00%	2.66%	-0.12%
2014/10/31	-0.83%	0.29%	-0.48%	1.01%	-2.16%	0.20%	-0.82%
2014/11/30	-1.30%	-6.97%	-1.50%	0.29%	-1.26%	-0.24%	-0.92%
2014/12/31	1.98%	0.42%	2.28%	3.24%	-1.33%	1.65%	1.98%
2015/01/31	-4.79%	1.76%	-2.23%	0.88%	-1.32%	1.81%	4.07%
2015/02/28	-2.37%	0.00%	1.15%	0.09%	-3.58%	0.16%	1.23%
2015/03/31	-0.60%	2.53%	1.92%	4.16%	3.39%	3.48%	-0.49%
2015/04/30	-1.00%	0.20%	-0.78%	-0.47%	-0.47%	-0.81%	1.24%
2015/05/31	3.13%	-1.30%	3.15%	-0.36%	1.69%	-2.09%	2.91%
2015/06/30	3.29%	0.40%	3.35%	2.74%	0.47%	2.67%	2.86%
2015/07/31	-0.69%	1.50%	1.24%	1.21%	-2.71%	1.52%	-1.05%
2015/08/31	4.96%	3.79%	3.91%	3.64%	1.96%	1.29%	2.11%
2015/09/30	6.04%	7.70%	3.62%	5.25%	2.09%	3.60%	4.79%
2015/10/31	-0.83%	-0.62%	-0.91%	-0.79%	1.26%	0.96%	-0.42%
2015/11/30	0.03%	2.37%	3.64%	4.50%	3.67%	2.88%	0.42%
2015/12/31	6.80%	6.22%	4.18%	5.58%	7.01%	4.93%	6.66%
2016/01/31	9.17%	12.33%	5.29%	9.28%	5.80%	8.09%	8.43%
2016/02/29	-1.67%	-1.01%	-4.45%	-3.82%	-2.05%	-5.15%	-2.57%
2016/03/31	-2.31%	-0.66%	-2.99%	-2.23%	2.51%	-0.43%	-1.42%
2016/04/30	-3.06%	-2.07%	-4.70%	-5.25%	-2.73%	-4.41%	-3.08%
2016/05/31	4.54%	5.25%	6.57%	4.87%	-0.11%	4.15%	3.40%
2016/06/30	-2.56%	1.34%	-4.09%	-2.01%	-0.56%	-2.52%	-1.09%
2016/07/31	-5.84%	-2.91%	-12.19%	-4.32%	-2.64%	-4.17%	-5.59%
2016/08/31	-3.50%	-2.11%	-5.16%	-4.83%	-3.52%	-4.48%	-3.61%
2016/09/30	2.14%	1.46%	2.55%	2.17%	1.69%	2.47%	1.71%
2016/10/31	-2.37%	-2.50%	-7.09%	-0.72%	-0.32%	-0.66%	-2.00%
2016/11/30	-2.19%	-4.26%	0.49%	-0.17%	-1.27%	-1.51%	-0.98%
2016/12/31	-2.92%	-7.58%	-0.03%	-0.61%	-3.00%	-0.95%	-2.89%
2017/01/31	-1.33%	-1.18%	-3.39%	-2.00%	-0.61%	-2.32%	-0.95%
2017/02/28	-2.62%	-1.11%	-1.43%	-2.73%	-0.10%	-1.20%	-2.13%

Continued on next page

Table B.1 – continued from previous page

Date	R\Euro	R\Yen	R\Sterling	R\USD	R\ASD	R\Ruppee	R\Sfranc
2017/03/31	-1.59%	-1.90%	-3.21%	-1.95%	-2.59%	-0.24%	-2.09%
2017/04/30	4.28%	6.51%	6.29%	4.00%	3.00%	6.04%	4.22%
2017/05/31	1.61%	-3.33%	0.92%	-1.51%	-2.90%	-1.27%	0.00%
2017/06/30	-1.20%	-1.62%	-3.88%	-2.86%	-1.18%	-2.88%	-0.94%
2017/07/31	4.25%	0.40%	3.33%	1.87%	4.92%	1.83%	2.57%
2017/08/31	3.35%	3.10%	0.52%	0.69%	2.31%	1.44%	0.24%
2017/09/30	0.15%	-1.39%	1.74%	-0.71%	-0.03%	-1.40%	-0.41%
2017/10/31	2.67%	1.96%	3.38%	4.04%	1.67%	3.00%	2.02%
2017/11/30	2.74%	2.96%	3.07%	2.89%	0.76%	3.25%	1.83%
2017/12/31	-5.71%	-6.59%	-5.16%	-6.54%	-6.38%	-5.72%	-6.17%
2018/01/31	-4.85%	-6.03%	-4.89%	-7.75%	-3.75%	-6.67%	-5.04%
2018/02/28	-1.75%	-0.39%	-1.80%	-3.18%	-4.00%	-4.35%	-0.24%
2018/03/31	-0.05%	1.90%	-0.09%	0.10%	-1.33%	-0.88%	-1.21%
2018/04/30	1.69%	0.65%	3.03%	2.09%	1.12%	1.17%	-0.03%
2018/05/31	-0.31%	1.61%	-0.95%	3.61%	1.41%	0.70%	0.50%
2018/06/30	4.73%	5.67%	4.59%	5.95%	5.47%	5.53%	6.65%

Table B.2

Standardised Rand foreign currency exchange rate monthly returns for period July 2013 to June 2018.

Index	Date	R\Euro	R\Yen	R\Sterling	R\USD	R\ASD	R\Ruppee	R\Sfranc
1	2013/07/31	-0.763	-1.022	-0.956	-0.560	-1.550	-1.318	-0.970
2	2013/08/31	1.051	0.924	1.015	0.436	0.069	-1.183	1.197
3	2013/09/30	-0.327	-0.739	0.291	-0.449	0.617	-0.799	-0.428
4	2013/10/31	0.407	0.175	0.190	-0.354	0.664	0.866	0.510
5	2013/11/30	0.463	0.062	0.719	0.717	0.285	0.334	0.445
6	2013/12/31	0.907	-0.577	0.888	0.354	-0.804	0.852	1.128
7	2014/01/31	1.254	1.132	1.401	1.319	1.167	1.383	1.134
8	2014/02/28	0.302	0.745	0.363	0.167	0.838	0.165	0.626

Continued on next page

Table B.2 – continued from previous page

Index	Date	R\Euro	R\Yen	R\Sterling	R\USD	R\ASD	R\Ruppee	R\Sfranc
9	2014/03/31	-0.381	-0.763	-0.546	-0.818	-0.398	-0.110	-0.382
10	2014/04/30	-0.715	-0.698	-0.398	-0.720	0.259	-0.346	-0.837
11	2014/05/31	-0.718	-0.240	-0.269	-0.578	-0.614	0.050	-0.863
12	2014/06/30	0.444	0.577	0.761	0.663	1.177	0.540	0.523
13	2014/07/31	-0.241	-0.022	0.197	-0.181	0.005	-0.286	-0.211
14	2014/08/31	-0.609	-0.400	-0.661	-0.136	-0.330	-0.506	-0.618
15	2014/09/30	-0.217	-0.458	0.015	0.678	-0.033	0.798	-0.169
16	2014/10/31	-0.348	0.007	-0.192	0.166	-0.838	-0.007	-0.411
17	2014/11/30	-0.496	-2.035	-0.476	-0.056	-0.500	-0.149	-0.448
18	2014/12/31	0.542	0.041	0.574	0.847	-0.530	0.467	0.553
19	2015/01/31	-1.601	0.418	-0.678	0.127	-0.523	0.521	1.276
20	2015/02/28	-0.833	-0.076	0.259	-0.115	-1.366	-0.018	0.296
21	2015/03/31	-0.276	0.634	0.473	1.128	1.231	1.064	-0.299
22	2015/04/30	-0.402	-0.020	-0.277	-0.286	-0.206	-0.336	0.299
23	2015/05/31	0.905	-0.442	0.813	-0.253	0.595	-0.755	0.875
24	2015/06/30	0.957	0.037	0.872	0.696	0.144	0.800	0.860
25	2015/07/31	-0.302	0.345	0.285	0.226	-1.044	0.427	-0.492
26	2015/08/31	1.486	0.989	1.027	0.971	0.697	0.351	0.600
27	2015/09/30	1.827	2.090	0.947	1.464	0.747	1.104	1.523
28	2015/10/31	-0.347	-0.250	-0.312	-0.384	0.436	0.242	-0.274
29	2015/11/30	-0.073	0.590	0.951	1.233	1.334	0.871	0.016
30	2015/12/31	2.067	1.673	1.100	1.564	2.578	1.539	2.169
31	2016/01/31	2.818	3.391	1.409	2.698	2.127	2.572	2.780
32	2016/02/29	-0.613	-0.361	-1.294	-1.313	-0.795	-1.754	-1.017
33	2016/03/31	-0.815	-0.261	-0.889	-0.827	0.901	-0.213	-0.619
34	2016/04/30	-1.053	-0.659	-1.365	-1.751	-1.050	-1.511	-1.193
35	2016/05/31	1.351	1.399	1.764	1.347	-0.074	1.285	1.043
36	2016/06/30	-0.896	0.301	-1.195	-0.759	-0.241	-0.894	-0.506
37	2016/07/31	-1.934	-0.895	-3.443	-1.465	-1.016	-1.433	-2.057
38	2016/08/31	-1.191	-0.670	-1.493	-1.622	-1.344	-1.535	-1.373
Continued on next page								

Table B.2 – continued from previous page

Index	Date	R\Euro	R\Yen	R\Sterling	R\USD	R\ASD	R\Ruppee	R\Sfranc
39	2016/09/30	0.592	0.335	0.649	0.521	0.597	0.737	0.463
40	2016/10/31	-0.834	-0.778	-2.028	-0.363	-0.152	-0.289	-0.818
41	2016/11/30	-0.777	-1.273	0.076	-0.195	-0.505	-0.565	-0.469
42	2016/12/31	-1.009	-2.207	-0.067	-0.332	-1.149	-0.381	-1.128
43	2017/01/31	-0.506	-0.408	-1.001	-0.757	-0.260	-0.829	-0.456
44	2017/02/28	-0.912	-0.387	-0.456	-0.980	-0.070	-0.464	-0.864
45	2017/03/31	-0.588	-0.611	-0.951	-0.740	-0.999	-0.149	-0.849
46	2017/04/30	1.271	1.754	1.687	1.082	1.085	1.903	1.326
47	2017/05/31	0.426	-1.012	0.195	-0.605	-1.114	-0.485	-0.129
48	2017/06/30	-0.466	-0.532	-1.137	-1.019	-0.471	-1.012	-0.453
49	2017/07/31	1.259	0.038	0.866	0.429	1.799	0.527	0.757
50	2017/08/31	0.974	0.795	0.085	0.068	0.828	0.399	-0.047
51	2017/09/30	-0.037	-0.466	0.423	-0.360	-0.044	-0.528	-0.270
52	2017/10/31	0.762	0.475	0.879	1.093	0.591	0.909	0.569
53	2017/11/30	0.783	0.757	0.794	0.742	0.252	0.990	0.502
54	2017/12/31	-1.892	-1.930	-1.493	-2.145	-2.409	-1.941	-2.258
55	2018/01/31	-1.620	-1.772	-1.416	-2.515	-1.429	-2.252	-1.869
56	2018/02/28	-0.637	-0.187	-0.561	-1.119	-1.522	-1.492	-0.213
57	2018/03/31	-0.102	0.459	-0.085	-0.113	-0.527	-0.358	-0.548
58	2018/04/30	0.450	0.108	0.782	0.496	0.385	0.312	-0.140
59	2018/05/31	-0.183	0.377	-0.323	0.961	0.494	0.158	0.044
60	2018/06/30	1.414	1.519	1.215	1.677	2.004	1.734	2.167

APPENDIX C

R-CODE

This appendix lists all the program code used in the study. The programming language used in **R**.

C.1 Code for 95% VaR example

The following block of code is used to obtain biplots provided in section 2.3.

```
#####
#      PCA biplot construction for 95% VaR data
#####

library(UBbipl)
library(readxl)

#Function for traditional PCA biplot

PCAbipl(X = VAR95.data[,-1], colours = "green",
        offset = c(0.3, 0.3,0.3, 0.3),
        offset.m = c(-0.25, -0.25, -0.25, -0.25,
                    -0.25,-0.25, -0.25),
```

```

    pch.samples = 15, pch.sample.size = 1.1,
    pos = "Hor", pos.m = c(4,4,2,2,4,1,1),
    side.label = c("right","right","left",
                  "left","right","right","left"),
    reflect = "x", ax.name.size = 0.7,
    n.int = c(3,10,10,5,10,10,3))

#Function for correlation PCA biplot

PCAbipl(X = VAR95.data[,-1], colours = "green",
        offset = c(0.3, 0.3,0.3, 0.3),
        offset.m = c(-0.25, -0.25, -0.25, -0.25,
                    -0.25,-0.25, -0.25),
        pch.samples = 15, pch.sample.size = 1.1,
        pos = "Hor", pos.m = c(4,4,2,2,4,1,1),
        side.label = c("right","right","left",
                      "left","right","right","left"),
        reflect = "x",ax.name.size = 0.7,
        n.int = c(3,10,10,5,10,10,3),
        correlation.biplot = TRUE)

#Function for traditional PCA biplot with prediction for sample 16

PCApredict16 <- PCAbipl(X = VAR95.data[,-1], colours = "green",
                        offset = c(0.3, 0.3,0.3, 0.3),
                        offset.m = c(-0.25, -0.25, -0.25, -0.25,
                                      -0.25,-0.25, -0.25),
                        pch.samples = 15, pch.sample.size = 1.1,
                        pos = "Hor", pos.m = c(4,4,2,2,4,1,1),
                        side.label = c("right","right",
                                      "left","left",

```

C.2. CODE TO CONSTRUCT REFINED BIPLLOT ADJUSTED FROM UBBIPLADDENDA

```

                                "right","right","left"),
                                reflect = "x", ax.name.size = 0.7,
                                n.int = c(3,10,10,5,10,10,3),
                                predictions.sample = 16, ort.lty = 2)

#Print prediction estimate

PCApredict16$predictions

#Determine correlation matrix

VaRcorrelation<-cov2cor(cov(VAR95.data[,-1]))
VaRcorrelation

```

C.2 Code to construct Refined biplot adjusted from UBbipl

The following block of code is inserted in line 266 of the `PCAbipl` function in the **R** package `UBbipl` developed by le Roux and Lubbe (2013) to construct Refined biplots.

```

#####
# Code to find eigen-values and
# -vectors from extreme correlation matrix
#####

# If dec==0 then construct Traditional biplot

if(dec==0){
  svd.out <- svd(X)

```

C.2. CODE TO CONSTRUCT REFINED BIPLLOT ADJUSTED FROM UBBIPLADDENDA

```

V.mat <- svd.out$v
U.mat <- svd.out$u
Sigma.mat <- diag(svd.out$d)
eigval <- (svd.out$d)^2
cl<<-cor(X)
print(cor(X))
}

# If dec==2 then construct Refined biplot

if(dec==2){

  #Function finds extremal correlation matrix

  h<-copstruc(X,method = "tail",k=up.stats)

  # Find SVD of extremal correlaton matrix

  cl<<-as.matrix(h$R)
  svd.out <- svd(cl)
  V.mat <- -svd.out$v
  D.vals<-sqrt((n-1)*svd.out$d)
  Sigma.mat <- diag(D.vals)
  eigval <- (D.vals)^2
  print(cl)

  # Identify extreme observations

  Identify_tails<-function(Xmult,k,prec){

    Xmult<-scale(Xmult)

```

C.2. CODE TO CONSTRUCT REFINED BIPLLOT ADJUSTED FROM UBBIPLADDENDA

```

n <- dim(Xmult)[1]
d <- dim(Xmult)[2]

# number of pairs

dL <- d * (d - 1) / 2
dInd <- subsets(d)

RANKM <- apply(Xmult, 2, rank)
Rtau <- sin(rho_tau_est(Xmult)$tau * pi / 2)

thetavec <- theta_function(seq(0.05, pi /
                             2 - .05,
                             length = prec), 2.5)

thetavec[round(prec/2, 0)] <- pi / 4

x <- sqrt(2) * cos(thetavec)
y <- sqrt(2) * sin(thetavec)

track <- 0

temp<-matrix(rep(0,n*d*d*d*prec),
             ncol=d,nrow=n*d*d*prec)

for(p in 1:prec){

  for ( dim1 in 1:d ){

    for (dim2 in dim1:(d-1) ){

```

C.2. CODE TO CONSTRUCT REFINED BIPLLOT ADJUSTED FROM UBBIPLADDENDA

```

    for (nc in 1:n){

        if((RANKM[nc, dim1] > n - k * x[p]) &
            (RANKM[nc, dim2] > n - k * y[p])){

            track <<- track+1

            temp[nc*dim2*dim1*p,]<-RANKM[nc,]

        } else {

            track <<- track+0

        }

    }

}

ext<-unique(temp)

ext<-ext[2:dim(ext)[1],]

store<-rep(0,dim(ext)[1])

for(i in 1:dim(ext)[1]){

    store[i]<-match(ext[i,],RANKM)[1]

}

Vals<- list(index=store,coords=Xmult[store,])

```



```

    return(Vals)
  }
  # Store extreme samples and change marker type

  values<-Identify_tails(X,up.stats,31)

  X.new.samples<-X[values$index,]

  Extreme.vals<<-X[values$index,]

  Extreme.index<<-as.vector(values$index)

}

```

C.3 Code for biplot simulation engine

The following block of code represents the function for the biplot simulation engine.

```

#####
# Biplot simulation engine
#####

source("PCAbipl.R")

Simulate_biplot <- function(sims,n,dims,dist,Sigma,mu,myMvd,k){

  qual.norm<-rep(0,sims)
  qual.tail<-rep(0,sims)

```

```
pred.norm<-matrix(rep(0,sims*dims),nrow=sims,ncol=dims)
pred.tail<-matrix(rep(0,sims*dims),nrow=sims,ncol=dims)

ad.norm<-matrix(rep(0,sims*dims),nrow=sims,ncol=dims)
ad.tail<-matrix(rep(0,sims*dims),nrow=sims,ncol=dims)

Avg.samp.err.N<-matrix(rep(0,sims*dims),nrow=sims,
                        ncol=dims)
Avg.samp.err.T<-matrix(rep(0,sims*dims),nrow=sims,
                        ncol=dims)

Total.samp.err.N<-rep(0,sims)
Total.samp.err.T<-rep(0,sims)

Avg.Extsamp.err.N<-matrix(rep(0,sims*dims),nrow=sims,
                           ncol=dims)
Avg.Extsamp.err.T<-matrix(rep(0,sims*dims),nrow=sims,
                           ncol=dims)

Tot.Extsamp.err.N<-rep(0,sims)
Tot.Extsamp.err.T<-rep(0,sims)

for(i in 1:sims){

  if(dist=="norm"){
    Xmult <- mvrnorm(n, mu, Sigma)
    x <- scale(Xmult)
  }
}
```

```

if(dist=="cop"){

  Xmult<-rMvdc(1000,myMvd)
  x <- scale(Xmult)

}
pcan<-NewBipl(x,dec=0,Title = paste("Normal ",i),
              label=FALSE)
pcaT<-NewBipl(x,dec=2,up.stats=k,
              Title = paste("tail ",i),label=FALSE)

estN.x<-x%*%pcan$V[,1:2]%*%t(pcan$V[,1:2])
estT.x<-x%*%pcaT$V[,1:2]%*%t(pcaT$V[,1:2])

Avg.samp.err.N[i,1:dims]<-sqrt(apply((x-estN.x)^2,
                                   2,mean))
Avg.samp.err.T[i,1:dims]<-sqrt(apply((x-estT.x)^2,
                                   2,mean))

Total.samp.err.N[i]<-sum(Avg.samp.err.N[i,1:dims])
Total.samp.err.T[i]<-sum(Avg.samp.err.T[i,1:dims])

estEN.x<-x[Extreme.index,]%*%pcan$V[,1:2]%*%t(pcan$V[,1:2])
estET.x<-x[Extreme.index,]%*%pcaT$V[,1:2]%*%t(pcaT$V[,1:2])

Avg.Extsamp.err.N[i,1:dims]<-apply((Extreme.vals-estEN.x)^2,
                                   2,mean)
Avg.Extsamp.err.T[i,1:dims]<-apply((Extreme.vals-estET.x)^2,
                                   2,mean)

```

```

                                ,2,mean)

Tot.Extsamp.err.N[i]<-sum(Avg.Extsamp.err.N[i,1:dims])
Tot.Extsamp.err.T[i]<-sum(Avg.Extsamp.err.T[i,1:dims])

qual.norm[i]<-pcan$PCA.quality
qual.tail[i]<-pcaT$PCA.quality

pred.norm[i,1:dims]<-as.vector(pcan$predictivity)
pred.tail[i,1:dims]<-as.vector(pcaT$predictivity)

ad.norm[i,1:dims]<-as.vector(pcan$adequacy)
ad.tail[i,1:dims]<-as.vector(pcaT$adequacy)

}
Mean.overall.error.N <- apply(Avg.samp.err.N,2,mean)
SE.overall.error.N <- apply(Avg.samp.err.N,2,sd)/sqrt(sims)

Mean.overall.error.T <- apply(Avg.samp.err.T,2,mean)
SE.overall.error.T <- apply(Avg.samp.err.T,2,sd)/sqrt(sims)

Mean.Extreme.error.N <- apply(Avg.Extsamp.err.N,2,mean)
SE.Extreme.error.N <- apply(Avg.Extsamp.err.N,2,sd)/sqrt(sims)

Mean.Extreme.error.T <- apply(Avg.Extsamp.err.T,2,mean)
SE.Extreme.error.T <- apply(Avg.Extsamp.err.T,2,sd)/sqrt(sims)

output<-list("Mean normal PCA sample residual"=

```

```

        c(Mean.overall.error.N,mean(Total.samp.err.N,
                                   na.rm = TRUE)),
"SE normal PCA sample residual"=
        c(SE.overall.error.N,sd(Total.samp.err.N,
                                   na.rm = TRUE)/sqrt(sims)),
"Mean tail PCA sample residual"=
        c(Mean.overall.error.T,mean(Total.samp.err.T,
                                   na.rm = TRUE)),
"SE tail PCA sample residual"=
        c(SE.overall.error.T,sd(Total.samp.err.T,
                                   na.rm = TRUE)/sqrt(sims)),
"Mean Extreme normal PCA sample residual"=
        c(Mean.Extreme.error.N,mean(Tot.Extsamp.err.N,
                                   na.rm = TRUE)),
"SE Extreme normal PCA sample residual"=
        c(SE.Extreme.error.N,sd(Tot.Extsamp.err.N,
                                   na.rm = TRUE)/sqrt(sims)),
"Average Extreme tail PCA sample residual"=
        c(Mean.Extreme.error.T,mean(Tot.Extsamp.err.T,
                                   na.rm = TRUE)),
"SE Extreme tail PCA sample residual"=
        c(SE.Extreme.error.T,sd(Tot.Extsamp.err.T,
                                   na.rm = TRUE)/sqrt(sims)))
}

```

C.4 Code for application of Refined biplots

The following block of code is used to obtain results in Chapter 5 and 6.

```
#####
# Simulation results
#####

#####
# Guassian copual with gamma marginals
#####

myCop<-normalCopula(param = 0.5, dim = 5)

myMvd <- mvdc(copula=myCop, margins=c("gamma",
                                     "gamma", "gamma", "gamma",
                                     "gamma"),
              paramMargins=list(list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2)))

x<-rMvdc(5000,myMvd)
x<-scale(x,center = TRUE,scale=TRUE)
pairs.panels(x,cor=FALSE)

pcan<-NewBipl(x,dec=0,label=FALSE,pos="Hor")
pcaT<-NewBipl(x,dec=2,up.stats=20,label=FALSE,
              pos="Hor")

estN.x<-x%*%pcan$V[,1:2]%*%t(pcan$V[,1:2])
estT.x<-x%*%pcaT$V[,1:2]%*%t(pcaT$V[,1:2])
```

```

Avg.samp.err.N<-apply((x-estN.x)^2,2,mean)
Avg.samp.err.T<-apply((x-estT.x)^2,2,mean)

Avg.samp.err.N
Avg.samp.err.T

c(Avg.samp.err.N,sum(Avg.samp.err.N))
c(Avg.samp.err.T,sum(Avg.samp.err.T))

estEN.x<-x[Extreme.index,]%*%pcan$V[,1:2]%*%t(pcan$V[,1:2])
estET.x<-x[Extreme.index,]%*%pcaT$V[,1:2]%*%t(pcaT$V[,1:2])

Avg.Extsamp.err.N<-apply((Extreme.vals-estEN.x)^2,2,mean)
Avg.Extsamp.err.T<-apply((Extreme.vals-estET.x)^2,2,mean)

Avg.Extsamp.err.N
Avg.Extsamp.err.T

c(Avg.Extsamp.err.N,sum(Avg.Extsamp.err.N))
c(Avg.Extsamp.err.T,sum(Avg.Extsamp.err.T))

pcan$PCA.quality
pcaT$PCA.quality

pcan$predictivity
pcaT$predictivity

pcan$adequacy
pcaT$adequacy

```

```
#####
# Gaussian copual with gamma marginals p=5 Simulation
#####

myCop<-normalCopula(param = 0.5, dim = 5)

myMvd <- mvdc(copula=myCop, margins=c("gamma", "gamma",
                                     "gamma","gamma",
                                     "gamma"),
              paramMargins=list(list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2)))

Extresults1<-Simulate_bipl(sims=100,n=500,dims=5,dist="cop",
                           myMvd = myMvd ,k=80)

Extresults1

#####
# Gaussian copual with gamma marginals p=4 Simulation
#####

myCop<-normalCopula(param = 0.5, dim = 4)

myMvd <- mvdc(copula=myCop, margins=c("gamma", "gamma",
                                     "gamma","gamma"),
              paramMargins=list(list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2)))
```



```

                                list(shape=2, scale=2),
                                list(shape=2, scale=2)))
Extresults1<-Simulate_bipl(sims=100,n=500,dims=4,dist="cop",
                            myMvd = myMvd ,k=80)

Extresults1
#####
# Gaussian copual with gamma marginals p=7 Simulation
#####

myCop<-normalCopula(param = 0.5, dim = 7)
myMvd <- mvdc(copula=myCop, margins=c("gamma", "gamma",
                                     "gamma","gamma","gamma",
                                     "gamma","gamma"),
              paramMargins=list(list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2)))

Extresults1<-Simulate_bipl(sims=100,n=500,dims=7,
                            dist="cop",myMvd = myMvd ,k=80)

Extresults1

#####
# Gumbel copual with gamma marginals p=5
#####

myCop<-gumbelCopula(param = 1.5, dim = 5)

```

```

myMvd <- mvdc(copula=myCop, margins=c("gamma", "gamma",
                                     "gamma", "gamma",
                                     "gamma"),
             paramMargins=list(list(shape=2, scale=2),
                               list(shape=2, scale=2),
                               list(shape=2, scale=2),
                               list(shape=2, scale=2),
                               list(shape=2, scale=2)))

x<-rMvdc(5000,myMvd)
x<-scale(x,center = TRUE,scale=TRUE)
pairs.panels(x,cor=FALSE)

pcaN<-NewBiplot(x,dec=0,label=FALSE,pos="Hor")
pcaT<-NewBiplot(x,dec=2,up.stats=20,label=FALSE,
                pos="Hor")

estN.x<-x%*%pcaN$V[,1:2]%*%t(pcaN$V[,1:2])
estT.x<-x%*%pcaT$V[,1:2]%*%t(pcaT$V[,1:2])

Avg.samp.err.N<-apply((x-estN.x)^2,2,mean)
Avg.samp.err.T<-apply((x-estT.x)^2,2,mean)

Avg.samp.err.N
Avg.samp.err.T

c(Avg.samp.err.N,sum(Avg.samp.err.N))
c(Avg.samp.err.T,sum(Avg.samp.err.T))

estEN.x<-x[Extreme.index,]%*%pcaN$V[,1:2]%*%t(pcaN$V[,1:2])

```

```

estET.x<-x[Extreme.index,]*%pcaT$V[,1:2]*%t(pcaT$V[,1:2])

Avg.Extsamp.err.N<-apply((Extreme.vals-estEN.x)^2,2,mean)
Avg.Extsamp.err.T<-apply((Extreme.vals-estET.x)^2,2,mean)

Avg.Extsamp.err.N
Avg.Extsamp.err.T

c(Avg.Extsamp.err.N,sum(Avg.Extsamp.err.N))
c(Avg.Extsamp.err.T,sum(Avg.Extsamp.err.T))

pcan$PCA.quality
pcaT$PCA.quality

pcan$predictivity
pcaT$predictivity

pcan$adequacy
pcaT$adequacy

#####
# Gumbel copual with gamma marginals p=5 Simulation
#####

myCop<-gumbelCopula(param = 1.5, dim = 5)

myMvd <- mvdc(copula=myCop, margins=c("gamma", "gamma",
                                     "gamma", "gamma",
                                     "gamma"),

```

```

        paramMargins=list(list(shape=2, scale=2),
                           list(shape=2, scale=2),
                           list(shape=2, scale=2),
                           list(shape=2, scale=2),
                           list(shape=2, scale=2)))

Extresults2<-Simulate_bipl(sims=100,n=500,dims=5,
                           dist="cop",myMvd = myMvd ,k=80)

Extresults2

#####
# Gumbel copual with gamma marginals p=4 Simulation
#####

myCop<-gumbelCopula(param = 1.5, dim = 4)

myMvd <- mvdc(copula=myCop, margins=c("gamma", "gamma",
                                     "gamma","gamma"),
              paramMargins=list(list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2)))

Extresults2<-Simulate_bipl(sims=100,n=500,dims=4,dist="cop",
                           myMvd = myMvd ,k=80)

Extresults2

#####
# Gumbel copual with gamma marginals p=7 Simulation
#####

```

```

myCop<-gumbelCopula(param = 1.5, dim = 7)
myMvd <- mvdc(copula=myCop, margins=c("gamma", "gamma",
                                     "gamma","gamma",
                                     "gamma", "gamma",
                                     "gamma"),
              paramMargins=list(list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2),
                                list(shape=2, scale=2)))

Extresults2<-Simulate_bipl(sims=100,n=500,dims=7,dist="cop",
                           myMvd = myMvd ,k=80)

Extresults2

#####
# 4 variate gumbel copula with marginals
# gamma(2,1);beta(2,2);t(5);gumbel(0,2)
#####

myCop<-gumbelCopula(param = 1.5, dim = 4)

myMvd <- mvdc(copula=myCop, margins=c("gamma", "beta",
                                     "t","gumbel"),
              paramMargins=list(list(shape=2, scale=2),
                                list(shape1=2, shape2=2),
                                list(df=5),
                                list(alpha=0,scale=2)))

```

```

x<-rMvdc(500,myMvd)
x<-scale(x)
pairs.panels(x,cor=FALSE)

pcan<-NewBipl(x,dec=0,label=FALSE,pos="Hor")
pcaT<-NewBipl(x,dec=2,up.stats=20,label=FALSE,
              pos="Hor",reflect="y")

estN.x<-x%%pcan$V[,1:2]%%t(pcan$V[,1:2])
estT.x<-x%%pcaT$V[,1:2]%%t(pcaT$V[,1:2])

Avg.samp.err.N<-apply((x-estN.x)^2,2,mean)
Avg.samp.err.T<-apply((x-estT.x)^2,2,mean)

Avg.samp.err.N
Avg.samp.err.T

c(Avg.samp.err.N,sum(Avg.samp.err.N))
c(Avg.samp.err.T,sum(Avg.samp.err.T))

estEN.x<-x[Extreme.index,]%%pcan$V[,1:2]%%t(pcan$V[,1:2])
estET.x<-x[Extreme.index,]%%pcaT$V[,1:2]%%t(pcaT$V[,1:2])

Avg.Extsamp.err.N<-apply((Extreme.vals-estEN.x)^2,2,mean)
Avg.Extsamp.err.T<-apply((Extreme.vals-estET.x)^2,2,mean)

Avg.Extsamp.err.N
Avg.Extsamp.err.T

```

```

c(Avg.Extsamp.err.N,sum(Avg.Extsamp.err.N))
c(Avg.Extsamp.err.T,sum(Avg.Extsamp.err.T))

pcan$PCA.quality
pcaT$PCA.quality

pcan$predictivity
pcaT$predictivity

pcan$adequacy
pcaT$adequacy

#####
# 4 variate gumbel copula with marginals Simulation
# gamma(2,1);beta(2,2);t(5);gumbel(0,2)
#####

myCop<-gumbelCopula(param = 1.5, dim = 4)

myMvd <- mvdc(copula=myCop, margins=c("gamma", "beta",
                                     "t","gumbel"),
              paramMargins=list(list(shape=2, scale=2),
                                list(shape1=2, shape2=2),
                                list(df=5),list(alpha=0,scale=2)))

Extresults3<-Simulate_bipl(sims=100,n=5000,dims=4,dist="cop",
                           myMvd = myMvd ,k=20)

Extresults3

```

```
#####
# Financial Application Exchange rates
#####

Exchange_rates <- read_excel("D:/Work/2nd year/
                             M thesis/Extreme value biplots/
                             Exchange rates.xlsx")

rates<-as.data.frame(Exchange_rates)
x<-scale(rates[,-1])
pairs.panels(x,cor=FALSE)

pcan<-NewBipl(x,dec=0,ax.type="predictive",
              offset = c(0.3, 0.3,0.3, 0.3),
              side.label =c("right","right","left",
                             "left","right","right","left"),
              pos="Hor")

pcaT<-NewBipl(x,dec=2,up.stats = 10,ax.type="predictive",
              pos="Hor",offset = c(0.1, 0.3,0.3, 0.3),
              pos.m=c(4,4,3,4,4,4,1),
              offset.m = c(0.05,0.05,-0.1,0.05,0.05,0.05,-0.7),
              reflect="y")

pcan<-NewBipl(x,dec=0,predictions.sample = 31,ax.type="predictive",
              offset = c(0.3, 0.3,0.3, 0.3),
              side.label =c("right","right","left",
                             "left","right","right","left"),
              pos="Hor",ort.lty = 2)

pcaT<-NewBipl(x,dec=2,predictions.sample = 31,up.stats = 10,
```



```

        ax.type="predictive",pos="Hor",
        offset = c(0.1, 0.3,0.3, 0.3),
        pos.m=c(4,4,3,4,4,4,1),
        offset.m = c(0.05,0.05,-0.1,0.05,0.05,0.05,-0.7),
        reflect="y",ort.lty = 2)

estN.x<-x%%pcan$V[,1:2]%%t(pcan$V[,1:2])
estT.x<-x%%pcaT$V[,1:2]%%t(pcaT$V[,1:2])

Avg.samp.err.N<-apply((x-estN.x)^2,2,mean)
Avg.samp.err.T<-apply((x-estT.x)^2,2,mean)

c(Avg.samp.err.N,sum(Avg.samp.err.N))
c(Avg.samp.err.T,sum(Avg.samp.err.T))

estEN.x<-x[Extreme.index,]%%pcan$V[,1:2]%%t(pcan$V[,1:2])
estET.x<-x[Extreme.index,]%%pcaT$V[,1:2]%%t(pcaT$V[,1:2])

Avg.Extsamp.err.N<-apply((Extreme.vals-estEN.x)^2,2,mean)
Avg.Extsamp.err.T<-apply((Extreme.vals-estET.x)^2,2,mean)

c(Avg.Extsamp.err.N,sum(Avg.Extsamp.err.N))
c(Avg.Extsamp.err.T,sum(Avg.Extsamp.err.T))

pcan$PCA.quality
pcaT$PCA.quality

```

```
pcan$predictions
```

```
pcaT$predictions
```

```
t(x[31,])
```

```
pcan$predictivity
```

```
pcaT$predictivity
```

```
pcan$adequacy
```

```
pcaT$adequacy
```

LIST OF REFERENCES

- Abbott, E.A. (1884). *Flatland: a Romance of many Dimensions*. Seeley & Co, London.
- Brand, H. (2013). *PCA and CVA biplots: A study of their underlying theory and quality measures*. Master's thesis, Department of Statistics and Actuarial Science, University of Stellenbosch, Stellenbosch, South Africa.
Available at: <http://hdl.handle.net/10019.1/80363>
- Chautru, E. *et al.* (2015). Dimension reduction in multivariate extreme value analysis. *Electronic Journal of Statistics*, vol. 9, no. 1, pp. 383–418.
- de Carvalho, M. (2016). Statistics of extremes: Challenges and opportunities. *Handbook of Extreme Value Theory and its Applications to Finance and Insurance*.
- Embrechts, P., McNeil, A. and Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond*, vol. 1, pp. 176–223.
- Everitt, B. (1994). Exploring multivariate data graphically: a brief review with examples. *Journal of Applied Statistics*, vol. 21, no. 3, pp. 63–94.
- Gabriel, K.R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, vol. 58, no. 3, pp. 453–467.
- Gardner-Lubbe, S., Le Roux, N.J. and Gowers, J.C. (2008). Measures of fit in principal component and canonical variate analyses. *Journal of Applied Statistics*, vol. 35, no. 9, pp. 947–965.
- Gower, J. and Hand, D. (1996). *Biplots*. 1st edn. Chapman and Hall, London.

- Gower, J.C., Lubbe, S.G. and Le Roux, N.J. (2011). *Understanding biplots*. John Wiley & Sons.
- Haug, S., Klüppelberg, C. and Kuhn, G. (2015). Copula structure analysis based on extreme dependence. *Statistics and Its Interface*, vol. 8, no. 1, pp. 93–107.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into Principal Components. *The Journal of Educational Psychology*, vol. 24, pp. 417–441.
- Hult, H. and Lindskog, F. (2002). Multivariate extremes, aggregation and dependence in elliptical distributions. *Advances in Applied probability*, vol. 34, no. 3, pp. 587–608.
- Jogdeo, K. (1982). Concepts of dependence. *Encyclopedia of Statistical Sciences*, vol. 1, pp. 324–334.
- Jolliffe, I.T. (2002). *Principal Component Analysis*. 2nd edn. Springer-Verlag, New York. ISBN 0387954422.
- Kendall, M.G. (1938). A new measure of rank correlation. *Biometrika*, vol. 30, no. 1/2, pp. 81–93.
- Klüppelberg, C. and Kuhn, G. (2009). Copula structure analysis. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 71, no. 3, pp. 737–753.
- Klüppelberg, C., Kuhn, G., Peng, L. *et al.* (2007). Estimating the tail dependence function of an elliptical distribution. *Bernoulli*, vol. 13, no. 1, pp. 229–251.
- Klüppelberg, C. and Stelzer, R. (2014). Dealing with dependent risks. In: *Risk-A Multi-disciplinary Introduction*, pp. 241–277. Springer.
- le Roux, N. and Lubbe, S. (2013). *UBbipl: UNDERSTANDING BIPLOTS: DATA SETS AND FUNCTIONS*. R package version 3.0.4.
Available at: <http://www.wiley.com/go/biplots>
- Longin, F. and Solnik, B. (2001). Extreme correlation of international equity markets. *The Journal of Finance*, vol. 56, no. 2, pp. 649–676.

- Mahalanobis, P.C. (1936). On the generalized distance in statistics. National Institute of Science of India.
- Mai, J. and Scherer, M. (2014). *Financial engineering with copulas explained*. Springer.
- Matthias, S. and Jan-frederik, M. (2017). *Simulating copulas: stochastic models, sampling algorithms, and applications*, vol. 6. Imperial College Press, London.
- Meissner, G. (2013). *Correlation Risk Modeling and Management: An Applied Guide including the Basel III Correlation Framework-With Interactive Models in Excel/VBA*. John Wiley & Sons.
- Nelsen, R.B. (2007). *An introduction to copulas*. Springer Science & Business Media.
- Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, vol. 2, no. 11, pp. 559–572. ISSN 1941-5982.
- Salmon, F. (2012). The formula that killed wall street. *Significance*, vol. 9, no. 1, pp. 16–20.
- Schmidt, R. and Stadtmüller, U. (2006). Non-parametric estimation of tail dependence. *Scandinavian Journal of Statistics*, vol. 33, no. 2, pp. 307–335.
- Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, vol. 8, pp. 229–231.
- Stone, R. (1947). On the interdependence of blocks of transactions. *Supplement to the Journal of the Royal Statistical Society*, vol. 9, no. 1, pp. 1–45.
- Van Blerk, S. (2000). *Generalising biplots and its application in S-Plus*. Master's thesis, Department of Statistics and Actuarial Science, University of Stellenbosch, Stellenbosch, South Africa.