UNIVERSITEIT·STELLENBOSCH·UNIVERSITY

jou kennisvennoot • your knowledge partner

BiplotGUI

Interactive Biplots in R

AM LA GRANGE

Department of Statistics and Actuarial Science
Stellenbosch University
South Africa

Version 0.0-2

12 August 2008

DRAFT

Preface

This manual documents the features of version 0.0-2 of the BiplotGUI package, the first release on CRAN. BiplotGUI makes it easy for users to construct and interact with biplots in R.

Chapter 1 gives an overview of biplots and introduces the new package. The setup of the package and some known issues are also discussed. Chapter 2 introduces the main features of the package through the exploration of three data sets, while all the features of the graphical user interface (GUI) are documented in full in Chapter 3. An appendix summarises the history of the package.

While the statistical output provided by the BiplotGUI has been checked for accuracy, the GUI itself occassionally gives console warnings or even crashes. It's quick to simply restart the GUI. If the behaviour is replicable, please contact me at the address below with the details. Otherwise, visit `http://biplotgui.r-forge.r-project.org/` for more information, for help, to report bugs, or to contribute to the package.

The BiplotGUI package was written as part of my Masters degree at Stellenbosch University in South Africa. I would like to thank my two supervisors, Prof Niël le Roux and Dr Sugnet Gardner-Lubbe, for their continued support. I would also like to thank Prof Patrick Groenen, whose suggestions inadvertently gave rise to the project.

Anthony la Grange

Contents

List of Figures

List of Tables

List of Screenshots

List of Keyboard Shortcuts

A   Shows points determined by PCO.

B   Shows points determined by metric MDS (identity transformation).

C   Shows points determined by non-metric MDS (monotone regression).

D   Shows points determined by semi-metric MDS (monotone spline transformation).

0   No biplot axes.

1   Constructs a PCA biplot.

2   Constructs a covariance/correlation biplot.

3   Constructs a CVA biplot.

4   Shows regression biplot axes.

5   Shows Procrustes biplot axes.

6   Shows circular non-linear biplot axes.

Ctrl+N   Interpolates a new sample.

Ctrl+L   Clears all additional descriptors from the biplot.

F11   Shows the currently displayed biplot region in an external window.

F12   Shows the currently displayed biplot in 3D in an external window.

Ctrl++   Shows the next set of legend entries.

Ctrl+-   Shows the previous set of legend entries.

Ctrl+G   Allows many graphical parameters to be set for the different groups of points.

Ctrl+A   Allows many graphical parameters to be set for the different biplot axes.

Ctrl+R   Reverts all the graphical parameters to their default values.

Ctrl+S   Saves the currently displayed biplot region in the currently selected file format.

Ctrl+C   Copies the currently displayed biplot region to the clipboard.

Ctrl+P   Prints the currently displayed biplot region.

F1   Enables pop-up help messages for the components of the main GUI window.

List of Abbreviations

ANOVA   analysis of variance

AOD   analysis of distance

CCO   canonical correlation analysis

CV   canonical variate

CVA   canonical variate analysis

GDP   gross domestic product

GUI   graphical user interface

IM   iterative majorisation

MDS   multidimensional scaling

MRAE   mean relative absolute error

PC   principal component

PCA   principal component analysis

PCO   principal coordinates analysis

PPP   purchasing price parity

RAE   relative absolute error

RDA   redundancy analysis

RRR   reduced rank regression

Chapter 1

Introduction

After a brief overview of biplots and R in Section 1.1, the BiplotGUI package itself is introduced in Section 1.2. The setup of the package is described in Section 1.3. Some known issues are listed in Section 1.4.

Chapter 2 illustrates some of the most important features of the package through the exploration of three data sets. All the features of the package are documented in full in Chapter 3. A history of the package is given in Appendix A.

1.1 Overview

In this section, a brief overview of biplots, existing biplot software, and the statistical programming language and environment R, is given. The style is intentionally non-mathematical. However, detailed references are provided for those who wish to gain a fuller understanding of the underlying theory.

1.1.1 Biplots

The biplot was first introduced by Gabriel in 1971. Since then many types of biplots have been developed. One can broadly distinguish between two approaches, the traditional and the new.

The traditional approach

Gabriel's 1971 biplot is now most commonly known as the classical biplot. The classical biplot is used to represent the elements of a matrix graphically, usually in two dimensions, although it is also possible to draw a biplot in three dimensions. The representation consists of two sets of vectors. The first set is used to represent the rows of the matrix, while the second set is used to represent the columns of the matrix. The (ij)th element of the matrix is approximated by the inner product of the ith vector for rows and the jth vector for columns. The absolute value of this inner product may be visually appraised as the length of the ith vector for rows multiplied by the length of the projection of the jth vector for columns onto it. It may also be appraised as the length of the jth vector for columns multiplied by the length of the projection of the ith vector for rows onto it. If the angle between the two vectors is obtuse, the approximated value is taken to be negative, while zero matrix entries are represented by pairs of near-orthogonal vectors. For a two-dimensional biplot, the representation is exact if the matrix is of rank two at most, while for matrices of rank three or higher, the best rank-two matrix approximation

---

[1]Parts of this chapter and most of the next are adapted from work submitted for publication by La Grange, Le Roux and Gardner-Lubbe.

may be found, which in turn may be represented exactly in a two-dimensional biplot. The prefix *bi* in 'biplot' refers to the simultaneous representation of both the rows and the columns of the matrix, rather than the dimensionality of the representation. The approximation of matrix entries by inner products characterises biplots of the traditional approach, but other quantities may also be approximated within the same biplots.

When the matrix in question is a data matrix of samples by variables, it is the samples and the variables which are represented as vectors. To easily differentiate between the two sets of objects, the samples are often represented as points rather than vectors. Gabriel (1971, Section 3) focuses on two special cases of the classical biplot, which is not unique. In the first, the Pythagorean distances between the points approximate the Mahalanobis (1936) distances between the samples, and the variable variances, covariances and correlations are also approximated. Greenacre (1984) calls this biplot the covariance biplot or the correlation biplot if the variables are first standardised to have unit variances (Underhill, 1990). In the second special case, the points are the same as those in the scatter plot of the scores of the first two principal components (PCs), and it is the usual Pythagorean distances between the samples which are approximated. The visual appraisal of variances, covariances and correlations is no longer valid in this biplot, except in a broad sense. In the BiplotGUI package, as in Gower and Hand (1996), the biplot of the second special case is referred to as the principal component analysis (PCA) biplot. However, in the literature, covariance/correlation biplots are often called PCA biplots.

Besides data matrices, other matrices of interest in multivariate statistics may be biplotted. Underhill's (1990) coefficient of variation biplot approximates the coefficients of variation of variables. When the samples are grouped, the points in Gabriel's (1972) canonical variate analysis (CVA) biplot coincide with the first two canonical variates (CVs), bringing them into relation with the original variables. In canonical correlation analysis (CCO), two groups of variables are measured on a common set of samples. Biplots for CCO (Haber and Gabriel, 1976; ter Braak, 1990; Graffelman, 2005) approximate either the within-group or between-group correlations, or both. Biplots for reduced rank regression (RRR) (ter Braak, 1990; ter Braak and Looman, 1994), also known as redundancy analysis (RDA), are closely related to those of CCO. Although no longer data matrices which are decomposed, data matrices remain the point of departure.

Bi-additive biplots (Bradu and Gabriel, 1978) also fall within the traditional approach to biplots. These biplots are drawn of two-way tables of numerical data rather than data matrices to provide diagnostic tools for bi-additive models. Other, more complicated models have also had biplots developed for them, amongst others by Gower (1990).

A new approach

While the traditional approach to biplots is based on the inner product approximation of matrix entries, the new approach—first formulated by Gower (1995) and strongly advocated by Gower and Hand (1996)—sees the biplot as the multivariate analogue of the scatter plot. Here the vectors that represent the variables are extended through the display space to become biplot axes. The biplot axes are calibrated with convenient, representative markers. Under the new approach, the (ij)th matrix entry is approximated simply by the value on the jth biplot axis onto which the point representing the ith sample orthogonally projects. Biplot axes used in this way are said to be *predictive*. Predictive biplot axes do not in general coincide with *interpolative* biplot axes which are used to graphically position new samples within an existing representation. In order to represent more variables than there are dimensions in a biplot, biplot axes are not in general orthogonal as scatter plot axes are.

Under the new approach, various mechanisms exist with which to position the points in biplots, and similarly, various mechanisms exist with which to position the biplot axes, be

they predictive or interpolative. In fact, the choice of mechanisms define the types of biplots. For certain types of biplots, the mechanisms used are entirely dependent on one another. So, for example, many existing multivariate techniques provide incidental representations of the samples of a data matrix in a few dimensions. These techniques include PCA (Pearson, 1901; Hotelling, 1933), CVA (Hotelling, 1935, 1936) and the analysis of distance (AOD) (Gower and Krzanowski, 1999), a distance-based generalisation of the analysis of variance (ANOVA). For each technique, a matching mechanism provides the appropriate biplot axes to give the PCA biplot (Gower and Hand, 1996, Chapter 2), CVA biplot (Gower and Hand, 1996, Sections 5.2–5.4, 5.6) and the AOD biplot (Krzanowski, 2004; Gardner, Le Roux, Rypstra and Swart, 2005), respectively. Similar comments apply to the covariance/correlation biplot (Gardner, 2001, Section 2.3.2).

For some biplots, the two mechanisms do no lock in as closely. Scaling techniques such as principal coordinates analysis (PCO) (Schoenberg, 1935; Gower, 1966)—also known as classical scaling—and various forms of multidimensional scaling (MDS) (eg Cox and Cox, 2001; Borg and Groenen, 2005), exist precisely as a way of representing high-dimensional samples as points in a few dimensions. For these representations of the samples, approximate biplot axes may be provided for by either the regression method (Kruskal and Wish, 1978) or the Procrustes method to give the regression and Procrustes biplots respectively (Gower and Hand, 1996, Chapter 3). The different mechanisms are therefore interchangeable.

All the biplots considered up to this point have been linear: in the traditional approach the variables were represented as vectors, in the new approach the variables are represented as linear axes. Gower and Harding (1988) generalise the linear biplot axes of the new approach to calibrated, interpolative non-linear biplot trajectories. Gower (1995) introduces the predictive circular non-linear biplot, while Gower and Ngouenet (2005) introduce the predictive orthogonal non-linear biplot. The points in these biplots are determined by PCO. When the distance metric is Pythagorean, the non-linear biplots reduce to the PCA biplot. For convenience, non-linear biplot trajectories are also often simply referred to as biplot axes. These biplot axes, however, are not drawn through the entire display space as linear biplot axes are.

The BiplotGUI package constructs biplots of the new approach advocated by Gower and Hand (1996).

1.1.2   Applications

It is especially Yan and Kang's (2003) adaptation, the GGE biplot, which is often used in practice. This biplot has become a standard tool amongst crop scientists. A recent example of its use is in Singh, Huerta-Espino, Sharma, Joshi and Trethowan (2007). The biplots described earlier have also been applied widely. The following are some interesting examples.

Crop science Vanoli *et al.* (2006) use PCA biplots to study the properties of apples which have been stored for six months.

Decision theory Losa, van den Honert and Joubert (2001) use the covariance biplot as a tool for conflict resolution.

Ecology Botes, McGeoch and van Rensburg (2006) use RRR biplots to study elephant- and human-induced changes to dung beetle assemblages.

Food science Adhikari, Heymann and Huff (2003) use CVA biplots to study the textural characteristics of low-fat, full-fat and smoked cheeses.

Marine biology Munday, Jones and Caley (1997) use CCO biplots to study coral-dwelling gobies.

Medical science Hol, Molster and Gjerdet (2008) use a covariance biplot to help determine whether the galvanic combination of titanium and stainless steel should be avoided in surgical implants.

Meteorology Gabriel and Mather (1986) use a PCO biplot in an exploratory data analysis of the 1951–1982 summer rainfall around Nelspruit, South Africa. This period includes ten years of cloud seeding.

Sports science Dawkins (1989) uses correlation biplots to compare the athletic records of different nations.

The coefficient of variation biplot has been applied less often; Underhill (1990) gives an example. So too the bi-additive biplot, at least in its original incarnation, although there are many derivatives.

All the examples cited thusfar have been of biplots of the traditional approach. This is not surprising considering the dearth of software for biplots with calibrated axes under the new approach (see the next section). Nevertheless, a considerable number have been published. Le Roux and Gardner (2005), for example, showcase many examples of biplots with calibrated linear axes from fields such as archaeology, agriculture, antiques, education, financial management, mineralogy and process control. Other recently published examples include PCA/CVA biplots in chemistry (Alves, Cunha, Amaral, Pereira and Oliveira, 2005), an AOD biplot in morphology (Gardner and Le Roux, 2006), PCA biplots in mineralogy (Jemwa and Aldrich, 2006) and a CVA biplot in cephalometry (Naidoo, Harris, Swanevelder and Lombard, 2006).

1.1.3 Other biplot software

Many statistical packages can be used to produce at least the simplest of biplots of the traditional approach. These include the major statistical packages Minitab (Minitab Inc, 2007), SPSS (SPSS Inc, 2007), STATA (StataCorp LP, 2007), STATISTICA (StatSoft Inc, 2007) and various products from SAS (2008). However, functionality is often limited, and results hard to obtain. Greater functionality is provided by the three dedicated biplot programs XLS-Biplot (Udina, 2005a,b), GGEBiplot (Yan and Kang, 2006) and BiPlot (Lipkovich and Smith, 2002a,b). XLS-Biplot is based on XLisp-Stat (Tierney, 1990) and has many useful features including a related web-server which can be used to construct biplots online. GGEbiplot is aimed mainly at agronomists, crop scientists and geneticists. It supplements the book by Yan and Kang (2003). BiPlot is an add-on for Excel, and although therefore potentially widely useful, it unfortunately has some minor but serious shortcomings (see Udina, 2005b).

The Genstat package (VSN International Ltd, 2007) can be used to calculate the coordinates of the elements of a biplot. These can then be drawn using a procedure from an add-on library. Other packages, offering some traditional biplot functionality, include Manet (Hofmann, 2000), for Macintosh only, and ViSta (Young, 2001). Some packages are aimed at ecologists—brodgar (Highland Statistics Ltd, 2007) with R, Canoco (Plant Research International, 2002) with CanoDraw (Smilauer, 2003), MVSP (Kovach Computing Services, 2008) and PC-ORD (MjM Software Design, 2007)—while the Excel add-on BrandMap (WRC Research Systems Inc, 2007) is aimed at marketers. All the software mentioned are for purchase, except XLS-Biplot, BiPlot, Manet and ViSta which are available free of charge. So too is R.

1.1.4 R

R (R Development Core Team, 2008) is a free statistical programming language and environment capable of producing high-quality graphics. Initiated by Ihaka and Gentleman (1996), it has become 'the *de facto* standard for statistical computing' (Greenacre, 2007, p. 213). It is

an open-source implementation of the S programming language, available for download for all the major platforms from the R Project homepage at http://www.r-project.org. The R core is updated regularly with minor version revisions released roughly every six months. The current version (as of June 2008) is R 2.7.1. Updates are relatively painless. R is easily extensible: a large number of user-written packages is available for download from repositories such as CRAN and BioConductor. These repositories can be accessed via the R Project homepage. As R has increased in popularity, so too has the number of books devoted to it. Recent general-topic books on R include Braun and Murdoch (2007), Chambers (2007) and Spector (2008). The book by Murrel (2005) deals specifically with graphics in R. Many more resources are freely available from the R Project homepage.

As far as biplots are concerned, the biplot method in R can be used to produce two variations of Gabriel's (1971) classical biplot. The classical biplot is most similar to the covariance/correlation biplot described earlier. Packages with support for traditional biplots include ade4 (Dray and Dufour, 2007, 2008), ade4TkGUI (Thioulouse and Dray, 2007a,b) and vegan (Oksanen *et al.*, 2008). In addition, the calibrate package (Graffelman, 2007) can be used to physically calibrate both scatter plot and biplot axes as described by Graffelman and van Eeuwijk (2005).

As opposed to the many solutions for biplots of the traditional approach listed in this section and in the previous one, software for biplots of the new approach has not been readily available. To produce biplots of the new approach, users have had to do their own programming in a suitable environment (for example, Gardner, 2001). To many potential users, such a task represents a major obstacle.

1.2 A new package

The primary aim with the BiplotGUI package is to make it easy to construct biplots of the kind advocated by Gower and Hand (1996) – biplots in which samples are represented as points and variables are represented as calibrated axes. Furthermore it is important to have quality graphics, and for the user to be able to interact with the biplots. For the reasons mentioned, R is an ideal environment for the creation of such a package.

The main features of the package can be summarised as follows:

- The package has a graphical user interface (GUI), making it easy to use even to the those new to R.

- Different types of biplots are supported: PCA, covariance/correlation, CVA, regression, Procrustes and circular non-linear.

- The points in regression and Procrustes biplots are determined by a separate scaling technique. PCO and various forms of MDS are available. In circular non-linear biplots, the points are fixed by PCO.

- Various distance metrics and data transformations are available.

- New samples can be interpolated onto existing biplots, and sample group means can be incorporated.

- Additional descriptors can be shown: convex hulls, alpha-bags, point densities and classification regions.

- Various diagnostic graphs are easiily accessible: graphs of convergence; point, group and axis predictivities; and Shepard diagrams.

- The package is highly interactive with support for dynamic variable value prediction, point and axis drag and drop, zooming, and the exploration of biplots in 3D.

- Biplots can be finely tailored. Amongst the elements that can be customised are titles, labels, legends, and more than 80 well-organised graphical parameters.

- Data can be imported from R objects or Excel files. Graphs can be saved to eight file formats. An array of context-specific numerical quantities are available for export back to R for further manipulation.

- The package is free and open-source.

Many of these features are illustrated in Chapter 2, and all are documented in Chapter 3.

1.3   Setup

R can be downloaded from CRAN via the R Project homepage at http://www.r-project.org/ for any of the three major platforms: Linux, MacOS X and Windows. BiplotGUI version 0.0-2 requires at least R 2.7.

At present, the package is intented to be run under Windows. To run, BiplotGUI requires a number of other packages to be installed.

The *essential* packages are

- MASS (Venables and Ripley, 2002)

- tcltk (R Development Core Team, 2008)

- tcltk2 (Grosjean, 2008)

- tkrplot (Tierney, 2008)

The first two of these should already be in place after a basic installation of R.

The *recommended* packages are

- KernSmooth (Wand, 2008), for the estimation of point densities

- rgl (Adler and Murdoch, 2008), for the creation of dynamic 3D biplots

- vcd (Meyer, Zeileis and Hornik, 2008), for the colours to be used in the representation of point densities

- xlsReadWrite (Suter, Treetron and Switzerland, 2006), for the import of data from Excel

BiplotGUI runs marginally better when the R console is set to SDI mode rather than the default MDI mode.

1.4   Known issues

- The graphs in the diagnostic tabs occassionally do not show when the tabs are opened. Click to another tab and back.

- Interpolative circular non-linear biplots do not always initialise correctly in 3D. If this happens, first display a 3D predictive circular non-linear biplot.

- Extensive use of the GUI leads to memory leaks. If the system becomes noticably slower, close the R console after saving, and re-open.

Chapter 2

Exploring

2.1 Introduction

In the next three sections, the most important features of the BiplotGUI package are illustrated through the exploration of three data sets. Further features are highlighted in Section 2.5. A systematic account of all features is given in Chapter 3.

2.2 Countries data

In this section, a country-comparative data set is used to show how the GUI is initialised, how its features are laid out, and how it can be used to explore multivariate data.

Table 2.1 gives measurements of eight variables for the countries with the 15 largest economies (by purchasing price parity (PPP) gross domestic product (GDP)) in 2007. These data have been derived largely from the 2007 CIA World Factbook (Agency, 2007) and are for illustrative purposes only. The countries (and their ranks) are Brazil (10), Canada (12), China (2), France (7), Germany (5), Italy (8), India (4), Indonesia (15), Japan (3), Mexico (13), Russia (9), South Korea (11), Spain (14), the UK (6), and the US (1). The variables are: PPP GDP per capita in US dollars (GDP); HIV/Aids prevalence as a percentage of the population (HIV.AIDS); life expectancy in years (LIFE.EXP.); military spending as a percentage of GDP (MIL.); oil consumption in barrels per annum per capita (OIL.CONS.); population in millions (POP.); number of fixed line telephones per capita (TEL.); and percentage unemployed (UNEMPL.). The aim is to represent these data in two or three dimensions so that a single, multivariate visual impression may be obtained, with the calibrated biplot axes incorporating information on the original variables.

2.2.1 Getting started

To load the BiplotGUI packge into R, the following command is entered at the R prompt, followed as usual by the enter key:

```
library(BiplotGUI)
```

It is assumed by this stage that the BiplotGUI package has already been set up as described in Section 1.3.

If the user is acquainted with R, the data to be passed to the GUI may be entered using the keybord or be imported into R and saved as a matrix or a data frame. The countries data have already been included in the package as a data frame, and may be viewed from within R by entering the commands

```
data(Countries)
Countries
```

Table 2.1. The countries data: eight variables measured on the countries with the 15 largest economies (PPP GDP) in 2007; countries listed in alphabetical order.

Country	GDP	HIV.AIDS	LIFE.EXP.	MIL.	OIL.CONS.	POP.	TEL.	UNEMPL.
Brazil	8710	0.7	72.2	2.6	4.0	190	204.2	9.6
Canada	35 370	0.3	80.3	1.1	25.1	33	622.3	6.4
China	7724	0.1	72.9	4.3	1.8	1322	278.4	4.2
France	29 852	0.4	80.6	2.6	11.3	64	543.5	8.7
Germany	31 941	0.1	79.0	1.5	11.7	82	657.8	7.1
India	3685	0.9	68.6	2.5	0.8	1130	44.0	7.8
Indonesia	4041	0.1	70.2	3.0	1.8	235	63.2	12.5
Italy	30 199	0.5	79.9	1.8	11.8	58	430.8	7.0
Japan	33 100	0.1	82.0	0.8	16.0	127	432.8	4.1
Mexico	10 570	0.3	75.6	0.5	6.6	109	182.7	3.2
Russia	12 350	1.1	65.9	2.7	6.5	141	283.6	6.6
S Korea	24 386	0.1	77.2	2.7	16.0	49	547.8	3.3
Spain	27 418	0.7	79.8	1.2	14.2	40	454.5	8.1
UK	31 723	0.2	78.7	2.4	11.0	61	552.9	2.9
USA	43 369	0.6	78.0	4.1	25.1	301	571.2	4.8

at the R prompt. To initialise the GUI with the countries data, the command

```
Biplots(Data = Countries)
```

is entered. Alternatively, a data set may be saved as an Excel 1997-2003 file, with its samples as rows and its variables as columns. If the countries data were saved in the working directory as an Excel file named 'Countries.xls', an equivalent call would have been

```
Biplots(excel = "Countries.xls")
```

Either of the previous two instructions launches the GUI and no further R commands are necessary for this data set.

2.2.2 The layout

Screenshot 2.1 shows the layout of the GUI after it has launched. Six regions are indicated:

- **The menu bar**, in addition to the settings box, contains the most important options. The three most important drop-down menus are Joint, Points and Axes. The biplots listed under Joint have both their points and axes determined according to a single, joint mechanism. The other biplots have their points determined from the Points menu and their axes determined from the Axes menu.

- **The biplot region** is where the biplot and optional title and legend are displayed. This space is responsive to mouse clicks and motion.

- **The settings box** may be used to set the action of the biplot axes, either predictive, centroid interpolative or vector sum interpolative. Various data transformations may be effected.

- **The diagnostics tabs** show output related to the currently displayed biplot. The convergence tab shows a graph of convergence; the points, groups and axes tabs show context-specific graphs of goodness-of-fit for points, groups and axes, respectively; the predictions tab shows dynamically predicted variable values; while the export tab allows various objects to be exported to R.

- **The kraal** is where points and axes may be kept, temporarily removing them from consideration.
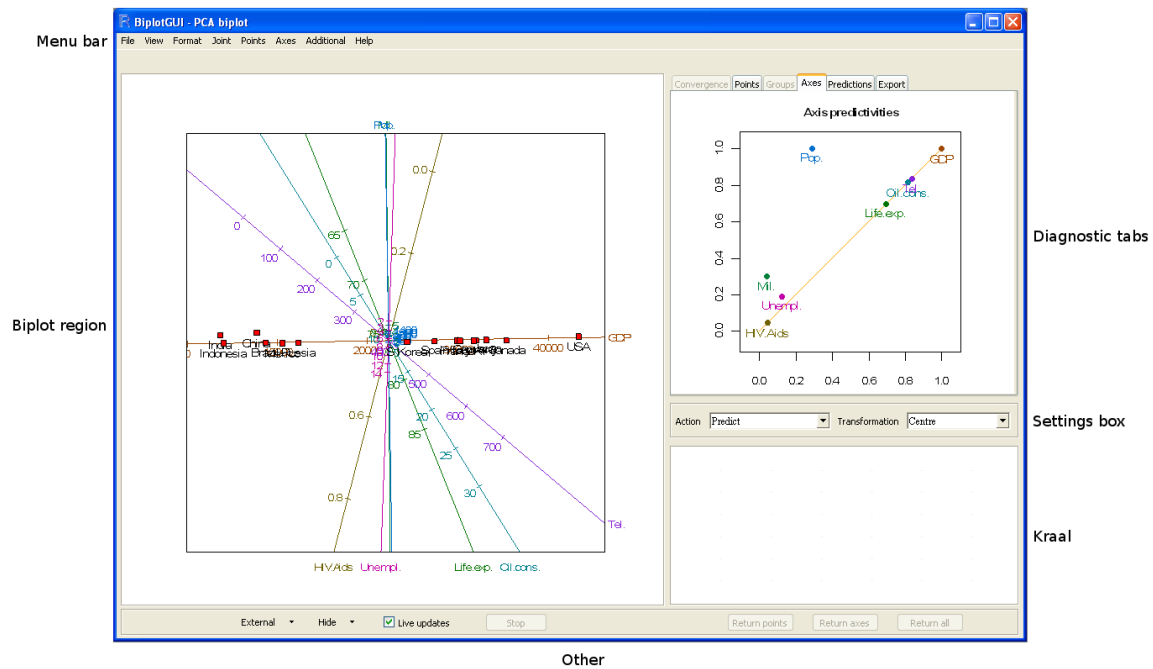
Screenshot 2.1: The BiplotGUI as it initially appears. A predictive PCA biplot of the countries data is shown towards the left; a graph of axis predictivities is shown towards the top right.

- **Other**. These options can be used to show the currently displayed biplot in an external window (in two or three dimensions), to control the biplot region or to control the kraal. While the GUI is busy, a progress bar is shown towards the left of this area.

Clicking Help → Show pop-up help activates pop-up help messages which appear as the mouse cursor is hovered over the components of the main GUI window.

2.2.3 Features

A PCA biplot with predictive axes is shown by default. For the countries data, this is the biplot shown towards the left of Screenshot 2.1. As should be the case for all biplots, a unit aspect ratio is used to ensure that distances within the biplot are properly represented. In this biplot, the points representing the countries lie ordered along a virtually straight line. In fact, the imagined line corresponds very closely to the biplot axis for GDP, and importantly, the line is almost horizontal. The reason for this becomes clear by looking at the GDP column of the countries data set in Table 2.1. The values of GDP are orders larger than those of the other variables. Therefore, that linear combination of the variables that has the largest possible variation (the first principal component) is heavily weighted towards GDP. In effect, GDP drowns out the other variables. To avoid this, the Centre, scale transformation is chosen from the available transformations in the settings box. This transformation independently transforms each variable to have unit variance and automatically updates the biplot to the one shown in Figure 2.1. Irrespective of the chosen transformation, biplot axes are always calibrated in terms of the original variable values. The first principal component in this new figure still ranks the countries from least to most wealthy, in some more complicated sense. The developed countries of the West, together with Japan and newly-industrialised South Korea, cluster in the south-east quadrant. Brazil, Russia and Indonesia lie more towards the west, with Mexico straddling the divide. India, and especially China, lie further away.
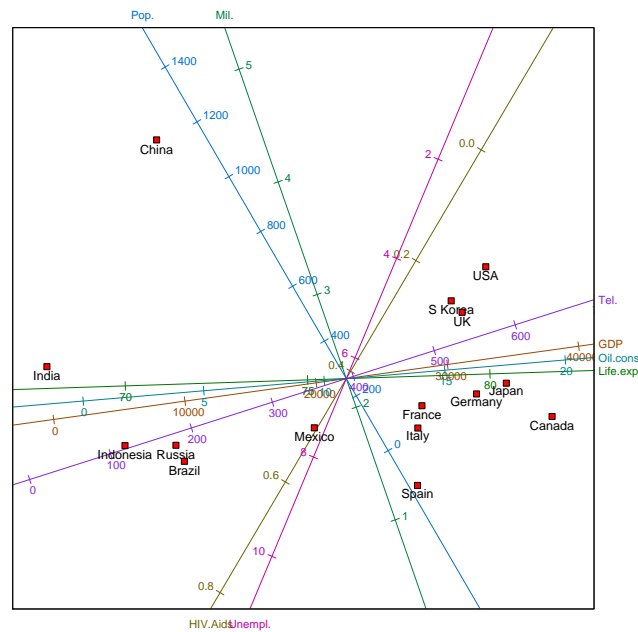
Figure 2.1. A predictive PCA biplot of the centred, scaled countries data.

While the relative positions of the points are interesting, biplots come into their own when the points are related to their original variable values through the axes. By right clicking inside the predictive linear biplot and selecting Predict cursor positions from the pop-up menu, an array of orthogonally projecting lines emanates from and follows the cursor as it moves over the biplot. If Predict points closest to cursor positions is selected instead, the lines project from the point closest to the cursor as it moves, rather than from the cursor itself. So for example, Figure 2.2 was created by hovering the cursor closer to the point for China than to any other point. These orthogonally projecting lines intersect the axes at the positions at which the optimal approximations to the original variables values are to be read off. It can be seen from Figure 2.2 that China scores relatively low on all the variables except population and military spending. As the cursor moves, these predictions are also given numerically, in real time, in the predictions tab. Dynamic prediction is disabled by right clicking inside the biplot and selecting Don't predict from the pop-up menu. (Some numerical predictions are given below.)

With many variables, a biplot may become crowded. A particular axis can be highlighted by right clicking it and selecting the Highlight option from the pop-up menu. Doing so greys the other axes and displays the true variable values of the highlighted axis above the corresponding points by default. The displays in the diagnostic tabs are shaded accordingly and orthogonal projections are drawn to the highlighted axis only. An example is shown in Figure 2.3, where GDP is highlighted and the variable value for China is predicted.

The question of course, is how good the biplot approximation is. This depends on both the points and the axes. As for the points, the 'quality' of the PCA approximation can be found by clicking to the export tab. For the countries data, 69.3% of the variation in the samples is accounted for by the first two principal components. Point and axis predictivities (Gower, Le Roux and Gardner-Lubbe, 2008) may also be calculated. Diagrams of these are available in the points and axes tabs, respectively. Those for the countries data are shown in Figure 2.4. These graphs were saved by right clicking them in the GUI and navigating the
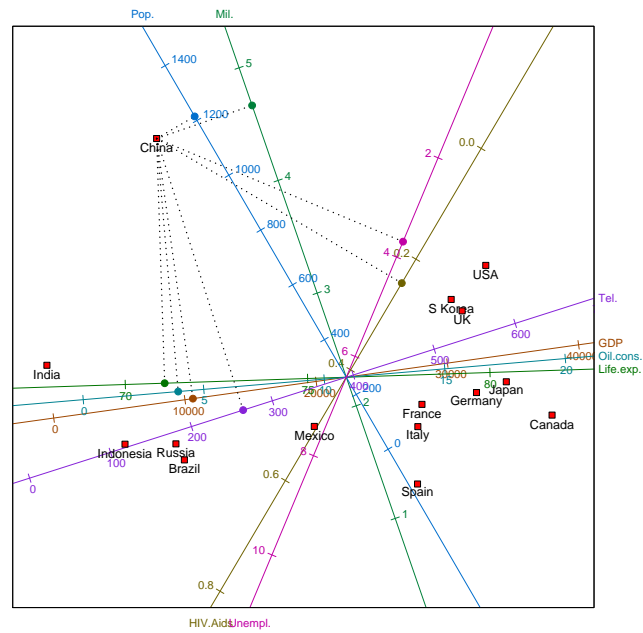
Figure 2.2. A predictive PCA biplot of the centred, scaled countries data with China projected onto all the biplot axes.
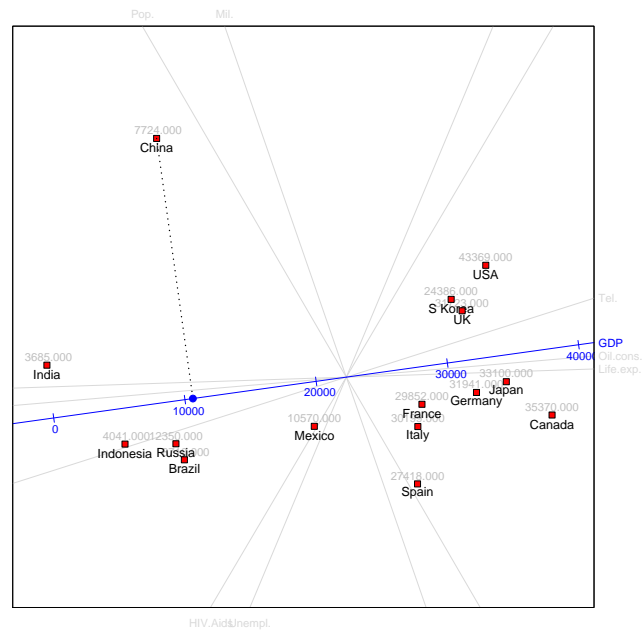
Figure 2.3. A predictive PCA biplot of the centred, scaled countries data with GDP highlighted and China projected.
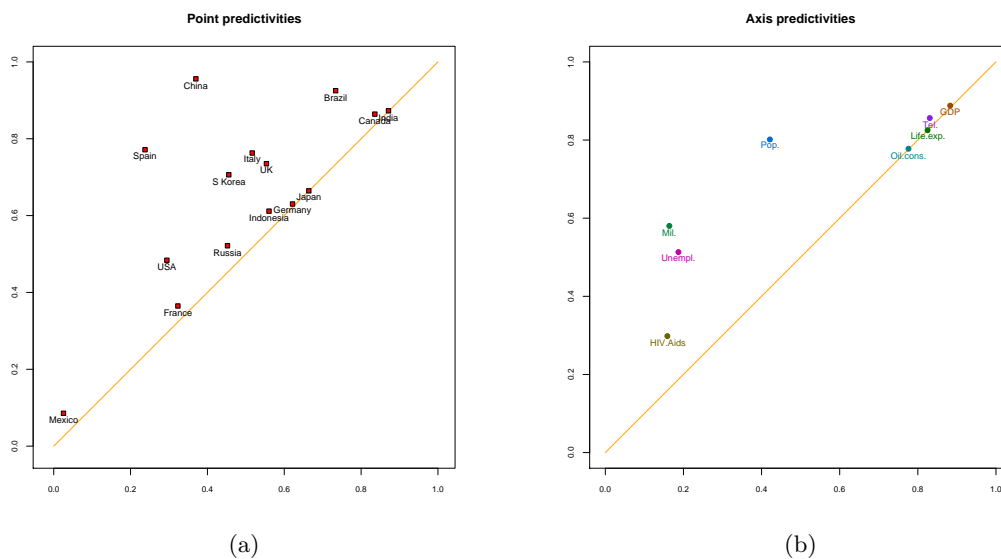
Figure 2.4. In (a), PCA point predictivities, and in (b), PCA axis predictivities, of the centred, scaled countries data.

Save as options of the pop-up menus. Points or axes further to the right in these diagrams are better represented in the first (or horizontal) biplot dimension. The closer to the top of the diagram, the better the point or axis is represented overall. India, Canada and Brazil are therefore relatively well represented in the first dimension. Japan, Germany and Indonesia are represented reasonably in the first dimension, but poorly in the second. France, the United States and Russia are poorly represented overall, and Mexico extremely poorly. China is the best represented country overall. The axis predictivities may be similarly interpreted. Predictivities are also available numerically from the export tab.

Another measure of the goodness of the approximation is relative absolute error (RAE). This measure can be calculated for any sample on any variable. The RAE is defined to be the absolute difference between the predicted value and actual value, expressed as a percentage of the range $(\max - \min)$ of the particular variable. For GDP, for example, the following output is obtained for the countries data from the export tab:

```
GDP
           Prediction  Actual  RelAbsErr%
Brazil         9330.3    8710         1.6
Canada        37282.7   35370         4.8
China         10606.7    7724         7.3
France        27669.1   29852         5.5
Germany       31869.4   31941         0.2
India            40.2    3685         9.2
Indonesia      5054.5    4041         2.6
Italy         27130.3   30199         7.7
Japan         34209.8   33100         2.8
Mexico        19392.0   10570        22.2
Russia         8865.5   12350         8.8
S Korea       30946.1   24386        16.5
Spain         26507.7   27418         2.3
UK            31644.7   31723         0.2
USA           33889.0   43369        23.9
```

Although the United States, Mexico and South Korea predict poorly on the GDP axis, the
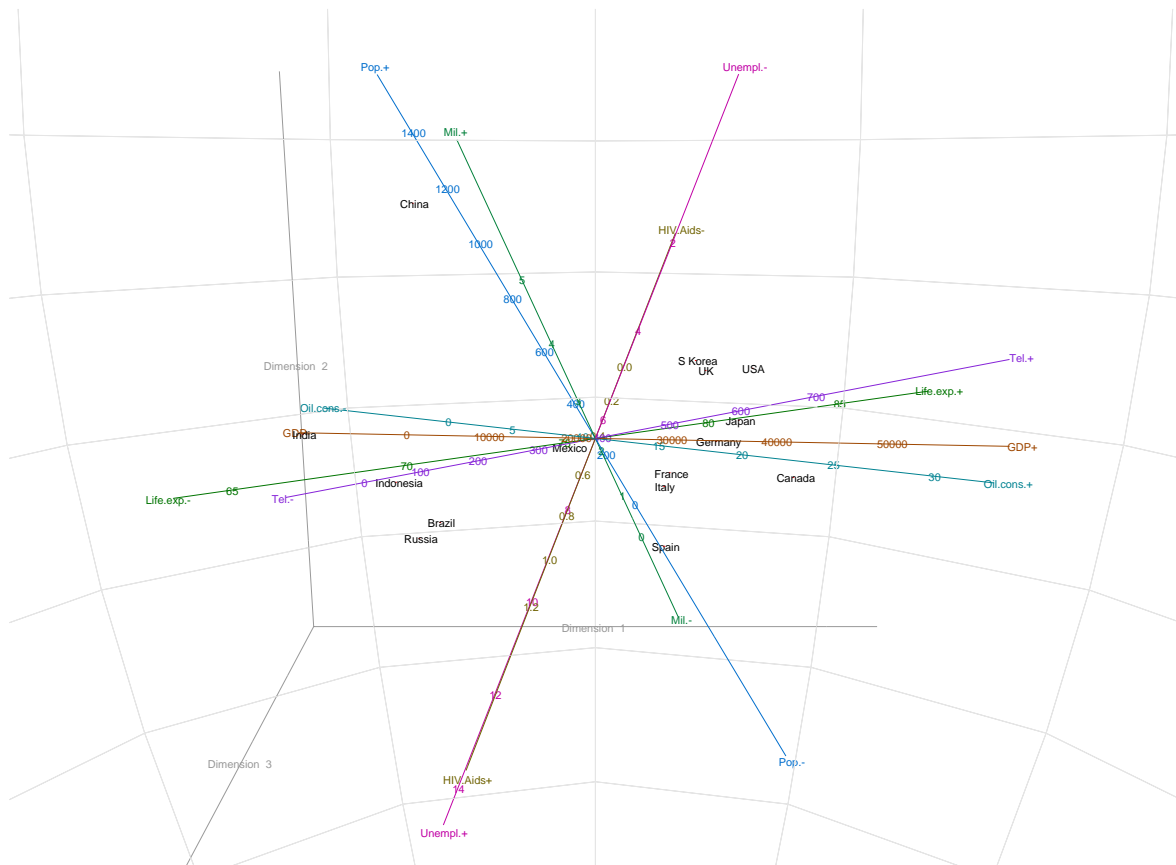
Figure 2.5. A predictive PCA biplot of the centred, scaled countries data, in three dimensions. This figure corresponds to the two-dimensional biplot in Figure 2.1.

overall configuration is optimal. By taking means over the samples, mean relative absolute errors (MRAEs) may be obtained for the different variables. From the export tab these are:

```
GDP  HIV.Aids Life.exp.    Mil. Oil.cons.     Pop.    Tel.  Unempl.
7.7     20.1      9.8     14.4    11.2      11.3     8.8    15.1
```

These error rates reinforce what is conveyed by the axis predictivities: that HIV/Aids prevalence, unemployment and military spending are relatively poorly represented, the other variables better. MRAEs are useful as a measure of the loss of information in biplots since they can be calculated for any type of biplot. Predictivities are defined only when certain orthogonal decompositions exist (Gower *et al.*, 2008), as they do in the case of PCA, CVA and AOD (Krzanowski, 2004; Gardner *et al.*, 2005) biplots.

For a biplot to be usable in printed form, it must necessarily be two-dimensional. However, assisted by a computer, a user may easily interact with a biplot in three dimensions. Three-dimensional, non-MDS biplots can be obtained by first clicking the External menu button at the bottom left of the GUI, followed by the In 3D option. Alternatively, the F12 shortcut key shown alongside the option can be used. Doing so renders the three-dimensional version of the currently displayed two-dimensional biplot in an external window. This feature makes use of the rgl package (Adler and Murdoch, 2008) and allows the biplot to be rotated and enlarged dynamically. Figure 2.5 shows the three-dimensional predictive PCA biplot of the countries data that corresponds to the two-dimensional version shown in Figure 2.1. A further 12.9% of the total variation in the samples is accounted for in the additional dimension. A

Figure 2.6. A predictive correlation biplot of the centred, scaled countries data.

360 degree 'fly-by' of the biplot is given by default as the window opens. This can be disabled by clicking File → Options.

A PCA approximation results from the projection of samples onto the plane of best fit. In a covariance biplot (Joint → Covariance/Correlation), these 'scores' are adjusted so that the cosines of the angles between the biplot axes approximate the correlations between the corresponding variables. The correlation biplot is the same as the covariance biplot, but with the variables first transformed to have unit variances (via the settings box in the usual manner). The correlation biplot of the countries data is shown in Figure 2.6. Three groups of variables are seen to be highly positively correlated (the angles between them are small):

- life expectancy, oil consumption, GDP, the number of telephone lines;

- HIV/Aids prevalence, unemployment;

- population, military spending.

Notice that the labels of the axes are attached to those ends of the axes that have the higher calibrations. This is the default option for all linear biplots. Alternatively, the axis labels may be given in a legend underneath the biplot, or no axis labels may be given whatsoever. These and similar options may be set via the View menu. Notice also that the option Joint → CVA is currently disabled, as canonical variate analysis requires the specification of group membership. This has not been provided in the case of the countries data, but will be for the antique furniture data in Section 2.3.

As opposed to dimension reduction by projection, in MDS the points are chosen so that the squared differences between the inter-sample *disparities* and the inter-point *distances* are explicitly minimised. The Points → MDS menu provides various options. These include taking the inter-sample disparities to be the inter-sample *dissimilarities* themselves (the identity transformation); retaining merely the order of the inter-sample dissimilarities by optimally
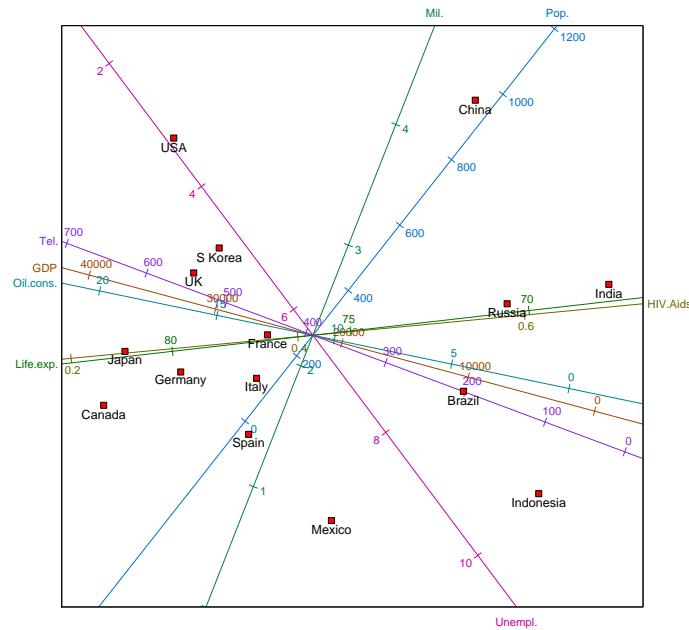
Figure 2.7. A predictive regression biplot of the centred, scaled countries data. The points are determined by metric MDS based on Pythagoras distance. The MDS representation is in terms of its principal axes.

transforming them into disparities (monotone regression, Kruskal, 1964); or monotonically smoothing the inter-sample dissimilarities into disparities (the monotone spline transformation, Ramsey, 1982, 1988). Therefore metric, non-metric and semi-metric MDS representations are available. The inter-sample dissimilarities are calculated according to the chosen distance metric. Four such metrics are currently available from the Points → Distance metric menu: Pythagoras distance, Square-root-of-Manhattan distance, Clark distance and Mahalanobis distance. Inter-point distances are always Pythagorean. An iterative majorisation (IM) algorithm (De Leeuw, 1977; De Leeuw and Heiser, 1980) is used to find the MDS solutions; for the implementations, see Borg and Groenen (2005, Chapters 8, 9). The IM algorithm converges uniformly, and usually leads to a local minimum, although in theory a saddle-point cannot be ruled out.

Figure 2.7 shows a metric MDS of the countries data, expressed in terms of its principal axes, with approximate regression biplot axes superimposed (from Points → MDS → In terms of principal axes, thereafter Points → MDS → Identity transformation). The default distance metric, Pythagoras distance, is retained. In this representation, the relative distances between the points are directly related to the corresponding dissimilarities between the countries. The United Kingdom and South Korea, therefore, are more similar to one another than they are to the other countries with respect to the eight variables. As the algorithm converges, updates of the configuration are shown in the biplot region, together with updates of the graphs in the diagnostic tabs. The Live updates option, however, may be disabled to increase the speed at which the algorithm converges. A graph of the stress values over iterations is given in the convergence tab. In this instance, from the export tab, convergence is reached after 96 iterations, with a final stress value of 45.8. A Shepard diagram (Borg and Groenen, 2005, Section 3.3; Shepard, 1962) can be found in the points tab and is shown in Figure 2.8. Each circle in the Shepard diagram represents a pair of samples. The horizontal axis indicates the
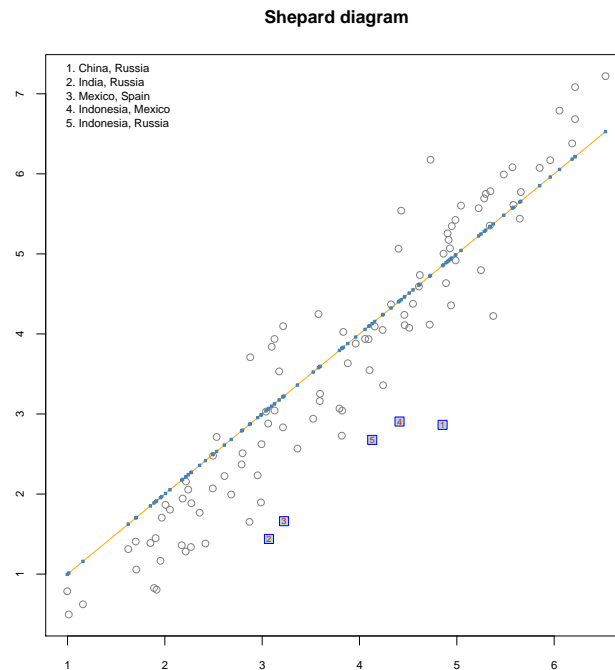
Figure 2.8. A Shepard diagram of the MDS of Figure 2.7.

inter-sample dissimilarity; the vertical axis indicates the corresponding inter-point distance. The blue dots on the yellow line (which generalises to a step function or a curve) indicate the disparities. Thus the closer the circles are to the line (or step function or curve), the better the overall fit. The five worst-fitting point pairs are identified in the top left corner of the diagram. The dissimilarity between China and Russia, therefore, is most poorly approximated by the points. The Points → MDS → Random initial configuration option forces the algorithm to start from a random configuration at each run. Otherwise, the last PCO or MDS solution is taken to be the new initial configuration, as was the case for Figure 2.7.

To conclude with the countries data, suppose that we feel that China is in many ways atypical and that we would like to see what the effect would be of removing it from consideration. To do so we need simply 'drag' the point representing China from the biplot into the kraal. We may also right click the point representing China and select Send to kraal from the pop-up menu. The biplot region is then automatically updated as if China were never part of the data set. The updated biplot is given in Figure 2.9. Russia's position relative to the other countries seems to have been most greatly affected. There has also been a re-alignment amongst the axes, most notably the axes for HIV/Aids, population and unemployment. Axes may also be removed to the kraal. Points and axes which have been removed to the kraal may be dragged back onto the biplot, or the kraal may be emptied of its points only, its axes only, or of both its points and axes simultaneously by making use of the buttons below it, or by right clicking inside it and selecting the desired option from the pop-up menu. At any stage, the points and/or axes of any representation may be hidden by clicking on the options in the hide menu button at the bottom of the window. Figure 2.10 is the same as Figure 2.9 but with the biplot axes hidden as described.
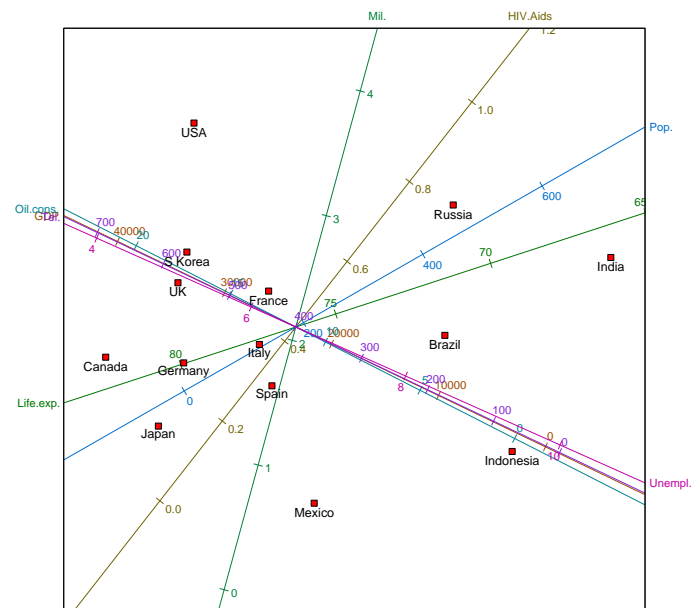
Figure 2.9. A predictive regression biplot of the centred, scaled countries data with China removed. The points are determined by metric MDS based on Pythagoras distance.
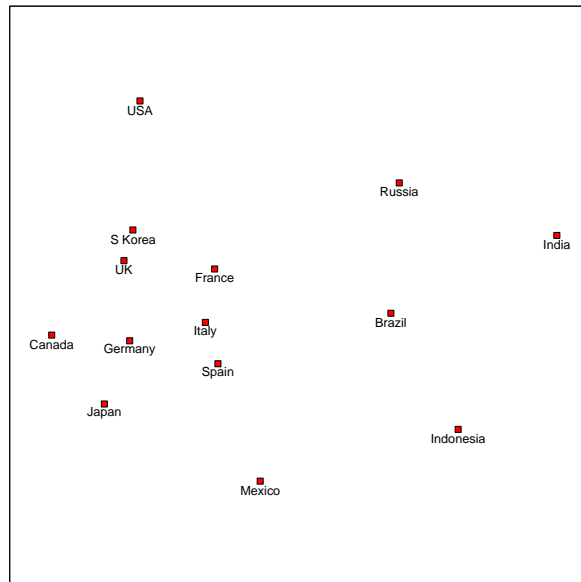

Figure 2.10. The same as Figure 2.9, but with the axes hidden: a metric MDS of the centred, scaled countries data with China removed; based on Pythagoras distance.

2.3   Antique furniture data

It is often of great interest to collectors, auctioneers and cultural historians to be able to correctly identify the type of wood used to make antique furniture. In the period between 1652 and 1900, wood from both the indigenous *Ocotea bullata* ('Stinkwood'), and the imported *Ocotea porosa* ('Imbuia'), were used to make Old-Cape furniture in South Africa. Being from the same genus and family (*Lauraceae*), it is often difficult to distinguish between the two types of wood based solely on a traditional analysis of colour, smell, and other observable characteristics. Burden, Gardner, Le Roux and Swart (2001) and Le Roux and Gardner (2005) make use of CVA biplots of anatomical measurements to distinguish between the species. A third species, *Ocotea kenyensis*, is also included in the analyses. The microscopically measured variables are: tangential vessel diameter in μm (VESD); vessel element length in μm (VESL); fibre length in μm (FIBL); ray height in μm (RAYH); ray width in μm (RAYW); and the number of vessels per mm^2 (NUMVES). The 37 observations are the mean values over fifty repeat-measurements made on 20 samples of *Ocotea bullata*, 10 of *Ocotea porosa*, and 7 of *Ocotea kenyensis*. The data are included in the BiplotGUI package as the data frame `AntiqueFurniture`, of which the first column contains the group specifications. The data may be viewed from within R by entering the following instructions at the R prompt of the R console:

```
data(AntiqueFurniture)
AntiqueFurniture
```

To initialise the GUI with the antique furniture data, the following command may be entered at the R prompt:

```
Biplots(Data = AntiqueFurniture[, -1], groups = AntiqueFurniture[, 1])
```

In other words, the data consist of all the columns of `AntiqueFurniture` except the first, while the group specifications are precisely the contents of the first column. Alternatively, if the data were saved as an Excel 1997–2003 file 'AntiqueFurniture.xls', with the group specifications given in, say, the first column, the call

```
Biplots(excel = "AntiqueFurniture.xls", ExcelGroupsCol = 1)
```

could have been made instead.

As was mentioned earlier, upon initialisation of the GUI, a predictive PCA biplot is shown by default. Clicking Joint → CVA produces a CVA biplot instead. This option has now become available since, in the call to the `Biplots` function, groups were defined. The predictive CVA biplot of the antique furniture data is shown in Figure 2.11. The positions of the points are determined by the first two CVs – those linear combinations of the original variables that maximally separate the group means, subject to certain restrictions (Krzanowski, 2000, Section 11.1). The group means themselves are shown as larger but corresponding symbols (activated by clicking Additional → Interpolate → Sample group means, retaining the default options). Since there is more than one group, an optional legend is included below the biplot by default. The mechanism for prediction is the same as before and is illustrated in the case of sample 29.

Figure 2.12 shows the same biplot, now overlain onto a two-dimensional density estimate of the points. The density estimate is obtained by clicking Additional → Point densities and accepting the default options (amongst other things, for the point densities to be estimated for all points, as opposed to certain groups of points only). The point densities are calculated using the default arguments to the `bkde2D` function of the KernSmooth package (Wand, 2008). Similar biplots can be found in Blasius, Eilers and Gower (2008).

Sometimes it is helpful to zoom into or out of portions of a biplot. This is done by right clicking on a focal point inside the biplot, and selecting the Zoom in or Zoom out option from the pop-up menu which appears. Figure 2.13 shows the CVA biplot of antique furniture,
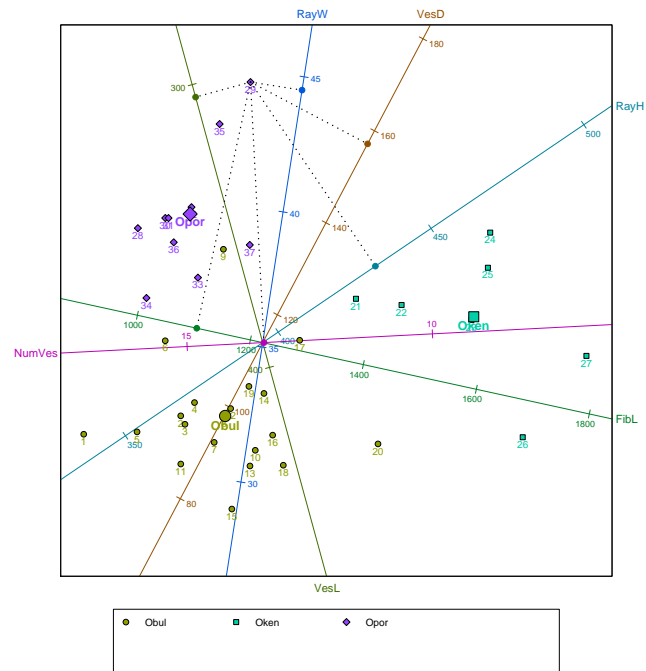
Figure 2.11. A predictive CVA biplot of the antique furniture data with sample 29 projected onto all the biplot axes. The group means are also shown.
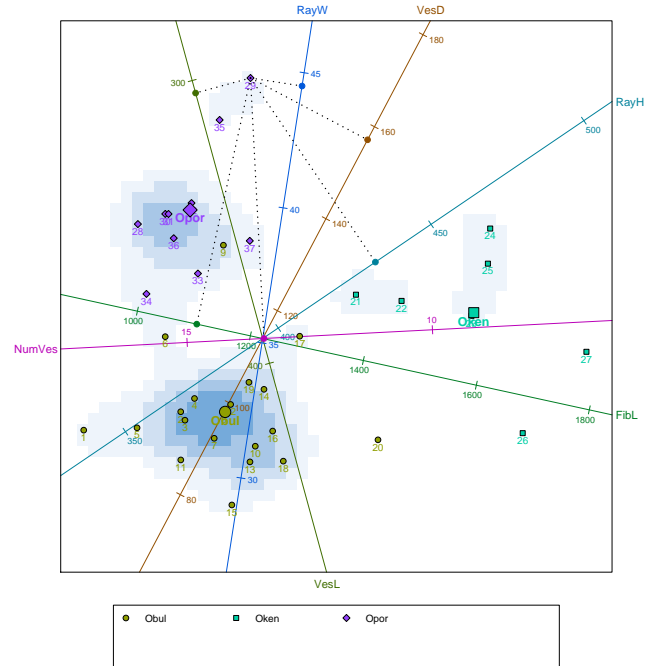

Figure 2.12. As in Figure 2.11, but with the biplot overlain onto a two-dimensional density estimate of the points.
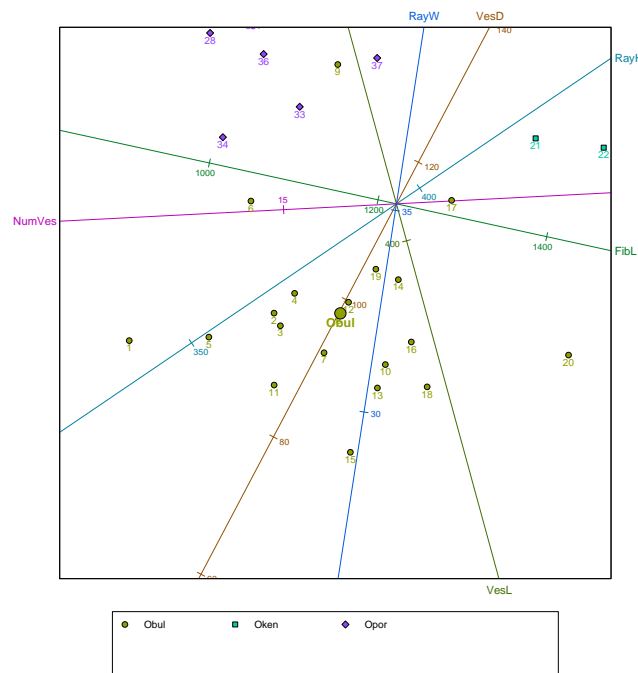
Figure 2.13. The same biplot as in Figures 2.11 and 2.12, but zoomed in around the mean of the species *Ocotea bulluta*.

enlarged around the mean of the species *Ocotea bullata*. The original view can be restored by clicking Reset zoom in the same pop-up menu.

To the left of Screenshot 2.2, a CVA biplot of the antique furniture data appears. From the settings box, it can be seen that the axes are not predictive; in fact they are vector sum interpolative. Also, the data have not been transformed, except for the obligatory centring of the columns to have zero means. In any case, CVA biplots are unaffected by the scaling of the variables to have unit variance.

Sample 18 has been dragged from the biplot in Screenshot 2.2 into the kraal. It has therefore not been taken into account in the construction of the biplot. However, using its variable values, it has subsequently been interpolated onto the biplot towards the bottom of the image (Additional → Interpolate → A New Sample). This is the most appropriate position for the sample in the existing biplot. It is reassuring that the positions assigned to sample 18 in Figure 2.11 and Screenshot 2.2 correspond so closely. This need not have been the case. Also notice that, notwithstanding the removal of sample 18, the calibrations and *directions* of the predictive and interpolative biplot axes differ. This is in general the case for CVA biplots.

The biplot in Screenshot 2.2 also sports colour-coded classification regions. These are the regions in the display space plane closest to the respective group means in a specified number of canonical dimensions, here the maximum number possible for the particular data set, six. The classification regions are included by selecting Classification regions from the Additional menu. They may be used for the classification of interpolated samples. For more on the links between biplots and discrimination, see Gardner and Le Roux (2005). Furthermore, by clicking Additional → Alpha-bags, alpha-bags (Gardner, 2001; Wurz, Le Roux, Gardner and Deacon, 2003) and Tukey medians have been superimposed for the species *Ocotea bullata* and *Ocotea porosa* (there are too few samples for an alpha-bag for *Ocotea kenyensis* to be constructed; with an appropriate warning, a convex-hull is displayed instead). Alpha-bags
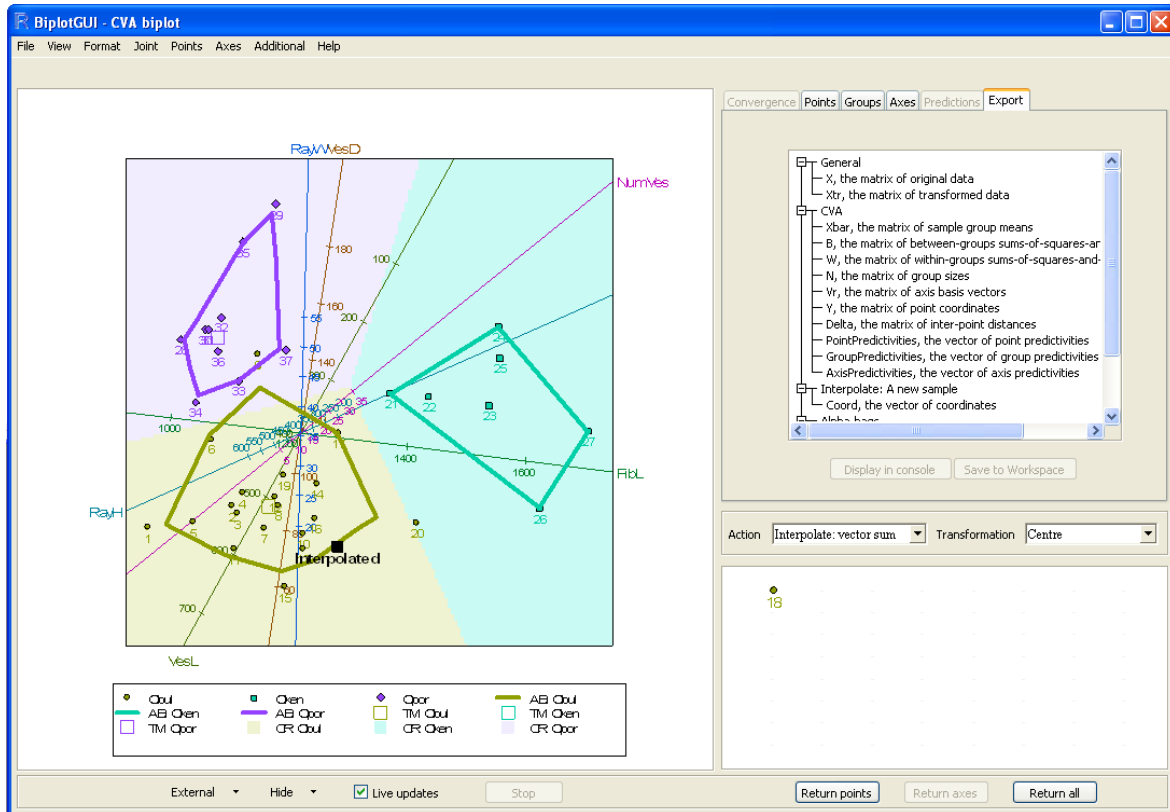
Screenshot 2.2: Towards the left, a vector sum interpolative CVA biplot of the antique furniture data with sample 18 removed to the kraal. Sample 18 has subsequently been interpolated onto the biplot to give its implied position. Classification regions are shown, as well as 90% alpha-bags for the species *Ocotea bullata* and *Ocotea porosa*. A convex hull surrounds the points of the species *Ocotea kenyensis*. The export tab is shown towards the top right.

are closely related to the bagplots of Rousseeuw, Ruts and Tukey (1999) and enclose regions that contain approximately the inner $100\alpha\%$ of samples, here 90% of the samples for the two species separately. The alpha-bags and convex hull do not overlap. This emphasises the high degree of separation between the species. For CVA biplots, group predictivities (Gower *et al.*, 2008) may also be calculated, in addition to the point and axis predictivities discussed earlier. A diagram of these is available in the groups tab. Finally, Screenshot 2.2 also shows the export tab. As has been explained, various objects are available for export from this tab, and these objects may be either displayed in the R console or be saved to the current R workspace. The list of available objects depends on what is shown in the biplot.

2.4 Fighter aircraft data

Measurements of four variables on 22 types of fighter aircraft were extracted by Cook and Weisberg (1982) from a report by Stanley and Miller (1979). Following Gower and Hand (1996), only the first 21 of these aircraft are included in the data and biplots below. The four variables are: specific power, proportional to power per unit weight (SPR); flight range factor (RGF); payload as a fraction of gross weight (PLF); and sustained load factor (SLF). These data can be found in the FighterAircraft data frame included in the BiplotGUI package. The GUI is initialised in the same way as it was for the countries data in Section 2.2.1.

Figure 2.14 shows a regression biplot of the fighter aircraft data with the points determined by PCO and the inter-sample dissimilarities calculated according to the Square-root-of-
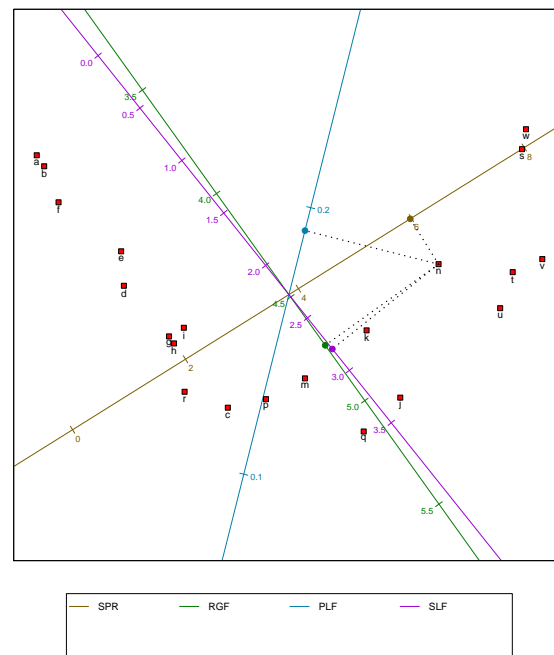
Figure 2.14. A predictive regression biplot of the fighter aircraft data, with points determined by PCO based on Square-root-of-Manhattan distance. The orthogonal prediction of the variable values of aircraft 'n' is shown.

Manhattan distance metric. The figure is obtained by clicking Points → Distance metric → Square-root-of-Manhattan and then Axes → Regression. Making use of orthogonal projection, the variable values for aircraft 'n' are predicted to be 7.342, 4.95, 0.202 and 3.25, respectively. These can be compared to the actual values, 5.855, 4.53, 0.172 and 2.50, respectively. Figure 2.15 shows the corresponding circular non-linear biplot (obtained by clicking Axes → Circular non-linear). Here prediction is performed by completing the circle which has, as diagonal, the line stretching from the origin of the biplot to the point to be predicted. The predicted values are read off at the points at which the circle intersects the axes (Gower and Hand, 1996, Section 6.3.2). If a particular axis is intersected at more than one position, the position closest to the point being predicted is used. If an axis isn't intersected at all, no prediction for the corresponding variable can be made. For aircraft 'n', the valid points of intersection are shown in the figure as small, filled circles on the circumference of the larger circle. From the predictions or export tabs, the circular non-linear predictions for aircraft 'n' are 6.090, 4.54, 0.174 and 2.90, respectively (these values depend on how finely the non-linear axes are constructed; by default 20 positions are taken into account from each calibrated marker to the next). Except for the fourth variable, the non-linear predictions are very close to the actual values. The non-linear biplot axes represent an appreciable improvement over the approximate linear axes in this case. This need not have been so.

2.5   Further features

There are two main ways in which graphs of the BiplotGUI can be customised. Basic customisation can be performed using the options of the View menu, while the Format menu options can be used to alter a large number of graphical parameters.
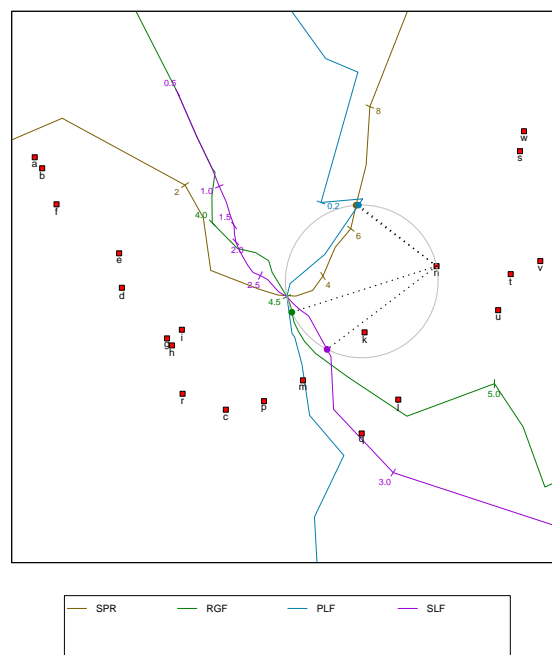
Figure 2.15. A predictive circular non-linear biplot of the fighter aircraft data, with points determined by PCO based on the Square-root-of-Manhattan distance metric. The circular prediction of the variable values of aircraft 'n' is shown.

Figure 2.16 shows the same predictive PCA biplot of the countries data as was shown in Figure 2.1. However, the biplot in Figure 2.16 has been modified by changing the default selections of the View menu. The View → Show title option places a title above the biplot; by default the title reflects the type of biplot, but it may be changed with the Format → Title option. Furthermore, the point labels have been removed by deselecting View → Show point labels. Instead of showing the axis labels around the edges of the biplot as in Figure 2.1, the labels in Figure 2.16 are shown in a legend (View → Show axis labels in legend). The View → Calibrate display space axes option calibrates the two dimensions of the biplot, although this is generally undesirable in biplots of the new approach (Gower and Hand, 1996, Section 2.6).

The Format menu allows virtually all the graphical parameters used internally by the package to be altered. The biplot in Figure 2.17 serves as an example. This biplot is the same as the one in Figure 2.11, but with some of the default graphical parameters changed. The Format → By group option allows the graphical parameters which relate to points, sample group means, convex hulls / alpha-bags and classification regions to be set for all groups simultaneously, or for a single group at a time. Screenshot 2.3 shows the dialogue box for this option as it appears when customising the points of the species *Ocotea bullata*. The parameter values appear as they have been set for Figure 2.17. The Format → Axes option similarly allows the graphical parameters for axes to be set, simultaneously for all axes or individually. Screenshot 2.4 shows the dialogue box of the graphical parameters for the axis RAYW, again with the parameters as they have been set for Figure 2.17. The graphical parameters used in variable value prediction and in the highlighting of axes can also be modified by clicking Format → Interaction, while diagnostic tab customisation may be performed by clicking Format → Diagnostic tabs. The Format → Reset all option reverts all the graphical parameters back to their default values. In all, more than 80 different graphical parameters may be set, often-times differently for different groups or axes. All these parameters are documented in
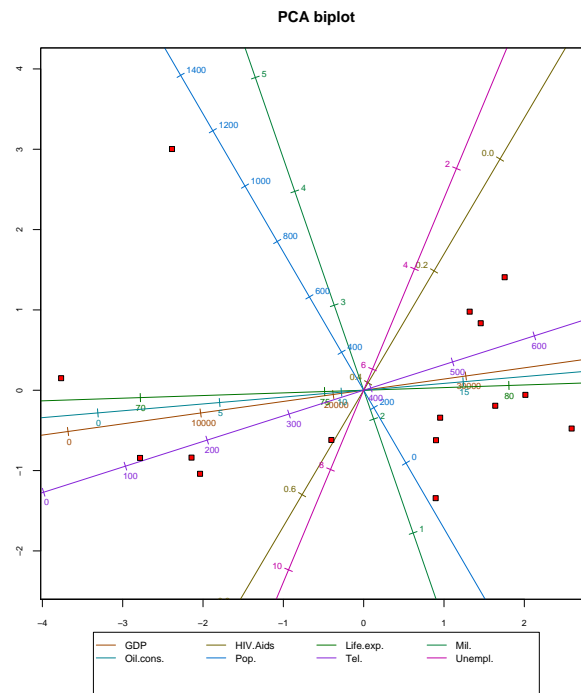
Figure 2.16. A modified version of the biplot given in Figure 2.1: a predictive PCA biplot of the countries data with a title, hidden point labels, axis labels in a legend, and the display space axes calibrated.

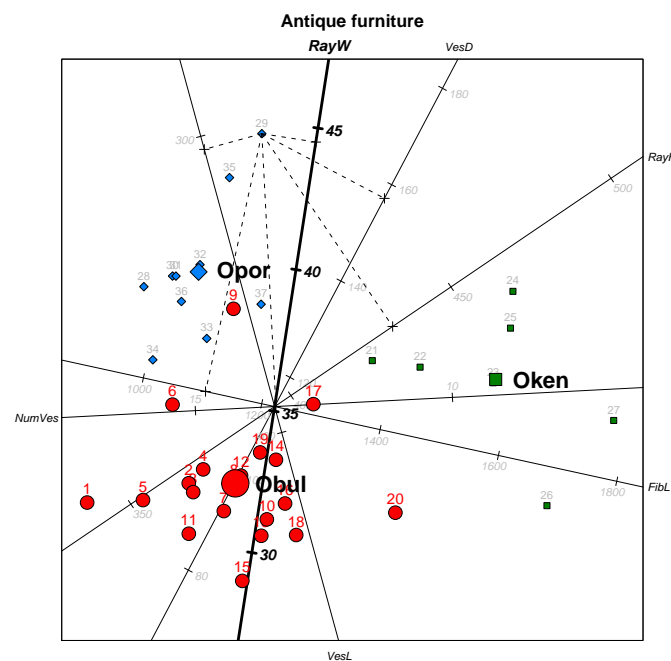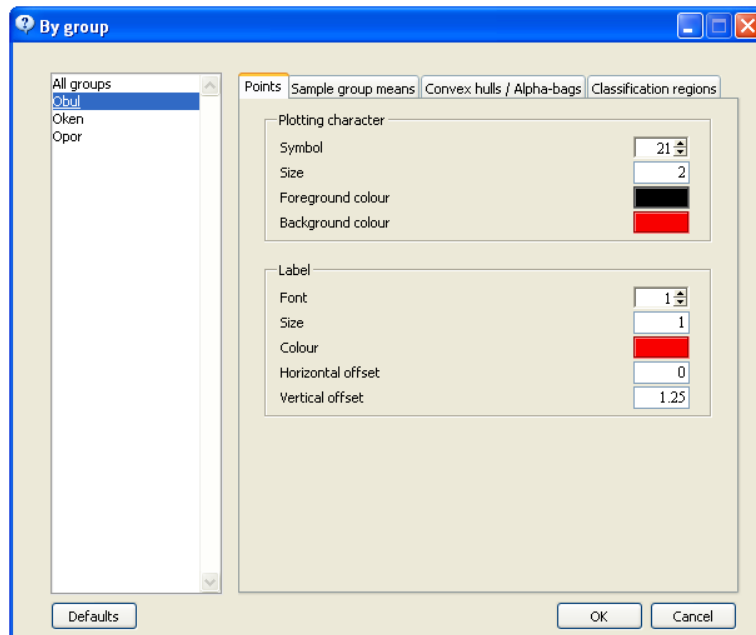

Figure 2.17. A modified version of the biplot given in Figure 2.11: a predictive CVA biplot of the antique furniture data with sample 29 projected onto all the biplot axes; the group means are shown.

Screenshot 2.3: The *Format → By group* dialogue box as it appears for the biplot in Figure 2.17.
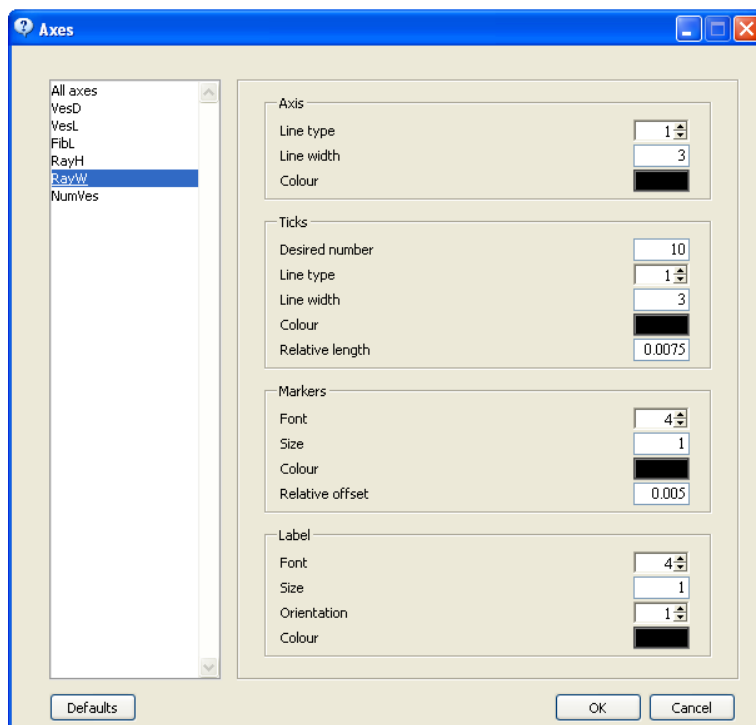
Screenshot 2.4: The *Format → Axes* dialogue box as it appears for the biplot in Figure 2.17.

detail in Chapter 3.

Naturally users would want to export the graphs produced using the BiplotGUI package. Biplot regions and diagnostic tab graphs can be saved in various file formats: PDF, Postscript, Metafile, Bmp, Png, Jpeg (50%, 75% and 100% quality) and PicTeX. Any graph can be saved by right clicking it and navigating the Save as menu. The biplot region may also be saved via the File → Save as menu. While the images shown onscreen are by necessity Metafile images, the images that appear in this article were saved in the PDF file format. Together with Postscript, such images are of the highest quality. Copy and print options are also available.

Chapter 3

Detail

The features of the GUI are documented in full in this chapter. For more information on the arguments of `Biplots`, the single function of the package, see the R help file for it by entering `?Biplots` at the R prompt of the R console.

3.1   General

While more than one instance of the GUI may be opened simultaneously from within the same R console (by issuing additional calls to `Biplots`), this often results in a console warning as the GUIs are closed. It is safe to use the console for other commands while running an instance of the GUI. GUIs can be run simultaneously from different instances of R.

3.2   List of Features

3.3   Features

After an introductory sentence, each feature is discussed under standard headings. These headings are listed below, together with a short description of each. Headings appear only when applicable.

DETAIL	More detailed information on the feature
DEFAULT VALUE	For graphical parameters, the default value
SHORTCUT	A keyboard shortcut to the feature
ALTERNATIVE TO	Mutually exclusive features
SREENSHOT	For dialogue boxes, a reference to a screenshot
REFERENCES	References to external material

Features are indented as in Section 3.2. The symbols used, from the outer-most to inner-most levels of indentation, are ▶, •, ▷, ◦ and ◇, respectively. Information on the options in dialogue boxes is given in framed sections.

▶ **Menu bar**
The primary interface of the GUI.

DETAIL · The menu bar is located at the top of the GUI. It consists of a number of drop-down menus, each with a number of entries. Entries may have sub-entries. The entries of menus are arranged around common themes of action. The Joint menu is for biplots with a joint mechanism for determining the positions of both points and axes. For other biplots, the Points menu determines the positions of the points while the Axes menu determines the positions

of the axes. Many entries contained in the menu bar can also be accessed from the pop-up menus which appear when various parts of graphs are right clicked.

SHORTCUTS · Often-used entries have dedicated keyboard shortcuts displayed to their right. These keyboard shortcuts are listed on page vi. Menu bar entries may also be accessed via the keyboard by pressing the Alt key together with the underlined letter in the name of the corresponding drop-down menu, followed by the sequence of underlined letters in the path to the entry.

SCREENSHOT · 2.1

- **File**
  Contains export and other options.

  DETAIL · The File menu lets the user save, copy or print the currently displayed biplot region. In addition, general options may be set, or the user may exit the GUI.

  ▷ **Save as**
    Lets the user save the currently displayed biplot region in a number of widely-used file formats.

    DETAIL · Depending on how the GUI has been resized onscreen, the proportions within a saved biplot region may differ from what is shown onscreen. To ensure consistency, saved biplot regions always have both their width and height set to 20.32 centimeters (8 inches). The PDF and Postscript file formats offer the highest quality, better even than that shown onscreen.

    SHORTCUT · Ctrl+S. Saves the currently displayed biplot region in the currently selected file format.

    ○ **PDF. . .**
      Saves the currently displayed biplot region as a PDF file.

      DETAIL · Together with Postscript, the PDF file format offers the highest quality.

      ALTERNATIVE TO · Postscript, Metafile, Bmp, Png, Jpeg 50% quality, Jpeg 75% quality, Jpeg 100% quality, PicTeX

      > A standard dialogue box appears which lets the user specify the name and location of the PDF file.

    ○ **Postscript. . .**
      Saves the currently displayed biplot region as a Postscript file (.ps).

      DETAIL · Together with PDF, the Postscript file format offers the highest quality.

      ALTERNATIVE TO · PDF, Metafile, Bmp, Png, Jpeg 50% quality, Jpeg 75% quality, Jpeg 100% quality, PicTeX

      > A standard dialogue box appears which lets the user specify the name and location of the Postscript file.

    ○ **Metafile. . .**
      Saves the currently displayed biplot region as a Windows Metafile (.wmf).

      ALTERNATIVE TO · PDF, Postscript, Bmp, Png, Jpeg 50% quality, Jpeg 75% quality, Jpeg 100% quality, PicTeX

> A standard dialogue box appears which lets the user specify the name and location of the Windows Metafile.

○ **Bmp...**

Saves the currently displayed biplot region as a bitmap (.bmp).

ALTERNATIVE TO · PDF, Postscript, Metafile, Png, Jpeg 50% quality, Jpeg 75% quality, Jpeg 100% quality, PicTeX

> A standard dialogue box appears which lets the user specify the name and location of the bitmap file.

○ **Png...**

Saves the currently displayed biplot region as a png file.

ALTERNATIVE TO · PDF, Postscript, Metafile, Bmp, Jpeg 50% quality, Jpeg 75% quality, Jpeg 100% quality, PicTeX

> A standard dialogue box appears which lets the user specify the name and location of the png file.

○ **Jpeg**

Lets the user save the currently displayed biplot region as a Jpeg file (.jpg, .jpeg).

◇ **50% quality...**

Saves the currently displayed biplot region as a Jpeg file at 50% quality.

ALTERNATIVE TO · PDF, Postscript, Metafile, Bmp, Png, Jpeg 75% quality, Jpeg 100% quality, PicTeX

> A standard dialogue box appears which lets the user specify the name and location of the 50% quality Jpeg file.

◇ **75% quality...**

Saves the currently displayed biplot region as a Jpeg file at 75% quality.

ALTERNATIVE TO · PDF, Postscript, Metafile, Bmp, Png, Jpeg 50% quality, Jpeg 100% quality, PicTeX

> A standard dialogue box appears which lets the user specify the name and location of the 75% quality Jpeg file.

◇ **100% quality...**

Saves the currently displayed biplot region as a Jpeg file at 100% quality.

ALTERNATIVE TO · PDF, Postscript, Metafile, Bmp, Png, Jpeg 50% quality, Jpeg 75% quality, PicTeX

> A standard dialogue box appears which lets the user specify the name and location of the 100% quality Jpeg file.

○ **PicTeX...**

Saves the currently displayed biplot region in the PicTeX format (.tex).

ALTERNATIVE TO · PDF, Postscript, Metafile, Bmp, Png, Jpeg 50% quality,
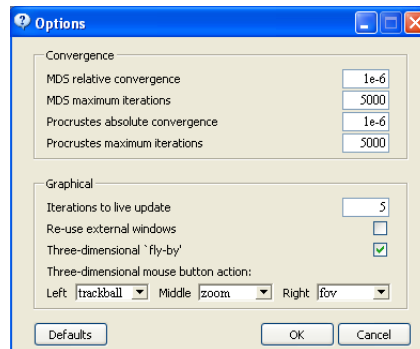
Screenshot 3.1: The *File → Options* dialogue box.

Jpeg 75% quality, Jpeg 100% quality

REFERENCES · See the R help file of the `pictex` function for the limitations of the conversion to this format.

> A standard dialogue box appears which lets the user specify the name and location of the PicTeX file.

▷ **Copy**
Copies the currently displayed biplot region to the clipboard as a Windows Metafile.

DETAIL · Depending on how the GUI has been resized onscreen, the proportions within a copied biplot region may differ from what is shown onscreen. To ensure consistency, copied biplot regions always have both their width and height set to 20.32 centimeters (8 inches). Biplots saved as PDF or Postscript files are of a higher quality than those copied to the clipboard.

SHORTCUT · Ctrl+C

▷ **Print. . .**
Prints the currently displayed biplot region.

DETAIL · Depending on how the GUI has been resized onscreen, the proportions within a printed biplot region may differ from what is shown onscreen. To ensure consistency, printed biplot regions always have both their width and height set to 20.32 centimeters (8 inches).

SHORTCUT · Ctrl+P

> A standard dialogue box appears which lets the user set various options before printing.

▷ **Options. . .**
Lets the user set various general options.

SCREENSHOT · 3.1

> **Convergence**
> Lets the user set various convergence criteria.
> *MDS relative convergence*
> If still running, the IM algorithm for MDS stops when the *proportional* decrease

in stress becomes smaller than this value.

DEFAULT VALUE · 10^{-6}

REFERENCES · Borg and Groenen (2005, pp. 191–192, 204–205)

MDS maximum iterations

If still running, the IM algorithm for MDS stops when the number of iterations reaches this value.

DEFAULT VALUE · 5000

REFERENCES · Borg and Groenen (2005, pp. 191–192, 204–205)

Procrustes absolute convergence

If still running, the Procrustes algorithm for interpolative biplots stops when the absolute decrease in stress first becomes smaller than this value.

DEFAULT VALUE · 10^{-6}

REFERENCES · Gower and Hand (1996, Section A.10.2), Gower and Dijksterhuis (2004, p. 57)

Procrustes maximum iterations

If still running, the Procrustes algorithm for interpolative biplots stops when the number of iterations reaches this value.

DEFAULT VALUE · 5000

REFERENCES · Gower and Hand (1996, Section A.10.2), Gower and Dijksterhuis (2004, p. 57)

Graphical

Lets the user set general graphical options.

Iterations to live update

The biplot region and relevant graphs in the diagnostic tabs are updated after this number of MDS iterations, and multiples of it.

DETAIL · The smaller the value of this option, the smoother the transitions in the graphs; the larger the value of this option, the smaller the time to convergence. Irrespective of the value of this option, the biplot region and the relevant graphs in the diagnostic tabs are only updated during MDS convergence if Live updates is checked below the biplot region.

DEFAULT VALUE · 5

Re-use external windows

Toggles between re-using windows for external graphics, or creating new windows.

DETAIL · If checked, the same window is used to display any two-dimensional external graphs that are produced. Such graphs can be created by clicking External → As is below the biplot region, or by right clicking the graphs of the diagnostic tabs and selecting External from the pop-up menu. If clear, new windows are used to show any additional two-dimensional graphs, retaining previous external graphs, if any. Similar comments hold for three-dimensional external biplots created by clicking External → In 3D below the biplot region. The windows of the two- and three-dimensional graphs operate independently.

DEFAULT VALUE · clear

Three-dimensional 'fly-by'

Toggles between including an initial 'fly-by' in three-dimensional biplots, or not.

DEFAULT VALUE · checked

> ***Three-dimensional mouse button action: left***
> The action of the left mouse button in three-dimensional biplots.
>
> DEFAULT VALUE · trackball
>
> REFERENCES · See the R help file of the `par3d` function of the `rgl` package (Adler and Murdoch, 2008) for details on the choices.
>
> ***Three-dimensional mouse button action: middle***
> The action of the middle mouse button in three-dimensional biplots.
>
> DEFAULT VALUE · zoom
>
> REFERENCES · See the R help file of the `par3d` function of the `rgl` package (Adler and Murdoch, 2008) for details on the choices.
>
> ***Three-dimensional mouse button action: right***
> The action of the right mouse button in three-dimensional biplots.
>
> DEFAULT VALUE · fov
>
> REFERENCES · See the R help file of the `par3d` function of the `rgl` package (Adler and Murdoch, 2008) for details on the choices.
>
> **Defaults**
> Reverts the options to their default values.
>
> **OK**
> Saves the options and returns to the GUI.
>
> **Cancel**
> Cancels the changes made since the dialogue box was opened, and returns to the GUI.

▷ **Exit**
Lets the user exit the GUI.

DETAIL · Upon exiting, all GUI settings changed during the session are lost. The user is returned to the R prompt of the R console.

> A standard dialogue box appears which asks the user to confirm whether or not to exit the GUI.

- **View**
  Lets the user customise the biplot region.

  ▷ **Show title**
  Shows a title above the biplot.

  DETAIL · The biplot retains the same size and relative position, irrespective of whether or not a title is shown.

  ▷ **Clip around points**
  Sets the biplot borders around the points, ignoring the axes.

  DETAIL · If selected, the biplot borders are set just large enough to show all the points and their labels in their entirety. Non-linear axes may not be visible in their entirety.

  ALTERNATIVE TO · Clip around points and axes

  ▷ **Clip around points and axes**
  Sets the biplot borders around the points and axes.

  DETAIL · If selected, the biplot borders are set large enough to show all the points, their labels, and the axes in their entirety. Available only for non-linear axes and

when the axes are not hidden.

ALTERNATIVE TO · Clip around points

▷ **Show point labels**
Shows the point labels in the biplot.
DETAIL · If selected, the biplot borders are enlarged so that all the point labels are visible in their entirety. Not available when the points are hidden.

▷ **Show point values**
Shows the variable values of a highlighted axis alongside the corresponding points in the biplot.
DETAIL · Only available when an axis has been highlighted.

▷ **Show group labels in legend**
Shows the group labels in the legend.
DETAIL · Not available when there is only one group of samples, or when the points are hidden.

▷ **Don't show axis labels**
Suppresses the axis labels.
DETAIL · Not available if there are no axes, or when the axes are hidden.
ALTERNATIVE TO · Show clinging axis labels, Show axis labels in legend

▷ **Show clinging axis labels**
Shows the axis labels at those edges of the axes with the higher calibrations.
DETAIL · Available only for linear axes that are not hidden.
ALTERNATIVE TO · Don't show axis labels, Show axis labels in legend

▷ **Show axis labels in legend**
Shows the axis labels in the legend.
DETAIL · Not available if there are no axes, or when the axes are hidden.
ALTERNATIVE TO · Don't show axis labels, Show clinging axis labels

▷ **Show Additional labels in legend**
Shows the labels of descriptors from the Additional menu in the legend.

▷ **Show next legend entries**
Shows the succeeding set of 16 legend entries, if any, in the legend.
DETAIL · The legend can contain up to a maximum of 16 entries at once. This option updates the legend to show the succeeding set of at most 16 entries. The option does not show the first set of entries after the last set. Not available when there are no succeeding entries, or when no legend is shown.
SHORTCUT · Ctrl++

▷ **Show previous legend entries**
Shows the preceding set of 16 legend entries, if any, in the legend.
DETAIL · The legend can contain up to a maximum of 16 entries at once. This option updates the legend to show the preceding set of 16 entries. The option does not show the last set of entries before the first set. Not available when there are no preceding entries, or when no legend is shown.
SHORTCUT · Ctrl+-

▷ **Calibrate display space axes**
Calibrates the two biplot dimensions.
DETAIL · This runs counter to the spirit of biplots of the new approach. Gower and Hand (1996) call such biplots 'bad'.
REFERENCES · Gower and Hand (1996, Section 2.6)
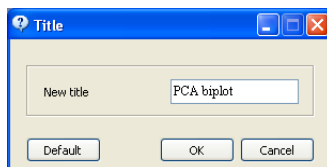
Screenshot 3.2: The *Format → Title* dialogue box.

- **Format**
  Lets the user customise the internally used graphical parameters.
  ▷ **Title...**
  Lets the user change the title.
  DETAIL · The title is shown only if View → Show title is selected.
  SCREENSHOT · 3.2

  > ***New title***
  > The new title.
  > DEFAULT VALUE · Depends on the currently displayed biplot; can be seen from the title bar of the GUI.
  > **Default**
  > Reverts the option to its default value.
  > **OK**
  > Saves the title, returns to the GUI, and redraws the biplot region.
  > **Cancel**
  > Cancels the changes made since the dialogue box was opened, and returns to the GUI.

  ▷ **By group...**
  Lets the user set graphical parameters for points, sample group means, convex hulls / alpha-bags and classification regions for a particular group, or for all groups simultaneously.
  DETAIL · When *All groups* is selected from the list to the left of the dialogue box, changes made to the parameter values to the right of the dialogue box affect all groups. Changes to particular groups can be made by clicking the group names in the list to the left of the dialogue box. Parameter values which differ amongst groups are left blank when *All groups* is selected.
  SHORTCUT · Ctrl+G
  SCREENSHOT · 3.3

  > *Only the Points tab is discussed here. The other tabs are discussed together with the options of the Additional menu.*
  > **Plotting character**
  > Lets the user change the graphical parameters of the points that represent the samples, by group.
  > ***Symbol***
  > Sets the symbol used to represent the points, by group.
  > DETAIL · Corresponds to R's `pch` argument. Possible values: `NA`, 0, ..., 25.
  > DEFAULT VALUE · If there is only one group, 22. Otherwise the values are recycled by group with the values 21 to 25.
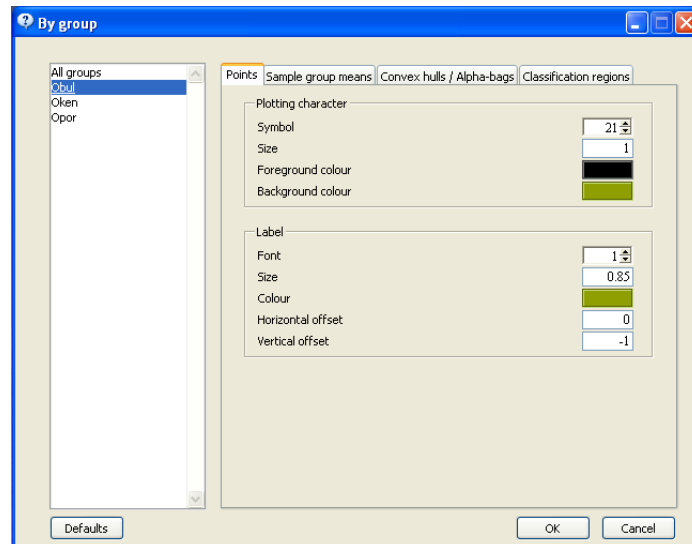
Screenshot 3.3: The *Points* tab of the *Format → By group* dialogue box.

Size

Sets the size of the points, by group.

DETAIL · Corresponds to R's `cex` argument. Larger values result in larger points. Possible values: positive real numbers.

DEFAULT VALUE · 1

Foreground colour

Sets the exterior colour of the points, by group.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "black"

Background colour

Sets the interior colour of the points, by group.

DETAIL · Corresponds to R's `bg` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · If there is only one group, "red". Otherwise equally spaced colours from the `hcl` spectrum.

Label

Lets the user change the graphical parameters of the point labels, by group.

Font

Sets the font of the point labels, by group.

DETAIL · Corresponds to R's `font` argument. Possible values: 1: regular; 2: bold; 3: italic; 4: bold italic.

DEFAULT VALUE · 1

Size

Sets the size of the point labels, by group.

DETAIL · Corresponds to R's cex argument. Larger values result in larger text. Possible values: positive real numbers.

DEFAULT VALUE · 0.85

Colour

Sets the colour of the point labels, by group.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · If there is only one group, "black". Otherwise equally spaced colours from the `hcl` spectrum.

Horizontal offset

Sets the horizontal offset of the point labels from the corresponding points, by group.

DETAIL · The point labels are shifted horizontally from the corresponding points with this multiple of the width of the letter 'x' (in the selected font size). Positive values shift the labels to the right; negative values shift the labels to the left.

DEFAULT VALUE · 0

Vertical offset

Sets the vertical offset of the point labels from the corresponding points, by group.

DETAIL · The point labels are shifted vertically from the corresponding points with this multiple of the height of the letter 'x' (in the selected font size). Positive values shift the labels towards the top; negative values shift the labels towards the bottom.

DEFAULT VALUE · −1

Defaults

Reverts the options of all the tabs of the dialogue box to their default values.

OK

Saves the options of all the tabs of the dialogue box, returns to the GUI, redraws the currently displayed biplot region and other graphs if necessary.

Cancel

Cancels the changes made since the dialogue box was opened, and returns to the GUI.

▷ **Axes...**

Lets the user set graphical parameters for a particular axis, or for all axes simultaneously.

DETAIL · When *All axes* is selected from the list to the left of the dialogue box, changes made to the parameter values to the right of the dialogue box affect all axes. Changes to particular axes can be made by clicking the variable names from the list to the left of the dialogue box. Parameter values which differ amongst axes are left blank when *All axes* is selected.

SHORTCUT · Ctrl+A

SCREENSHOT · 3.4

Axis

Lets the user change the graphical parameters associated with the axis that represents the variable.
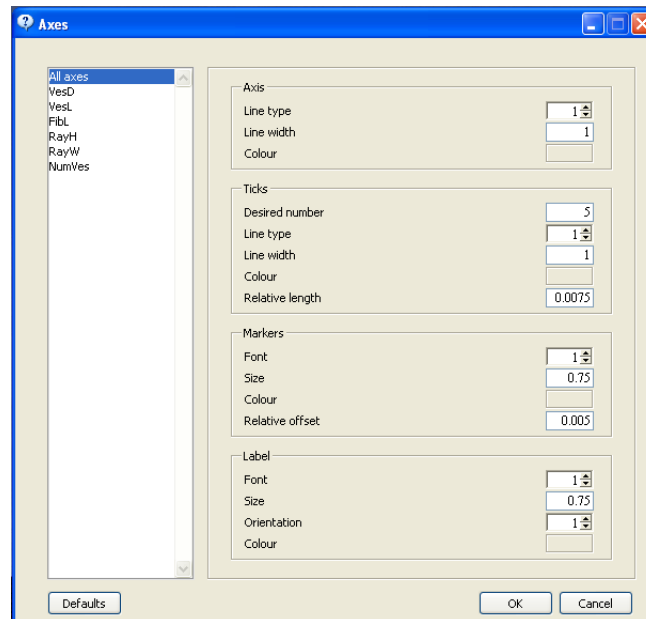
Screenshot 3.4: The *Format → Axes* dialogue box.

Line type

Sets the type of line used to draw the axis.

DETAIL · Corresponds to R's `lty` argument. Possible values: 0, . . . , 6.

DEFAULT VALUE · 1

Line width

Sets the width of the axis.

DETAIL · Corresponds to R's `lwd` argument. Larger values result in wider axes. Possible values: positive real numbers.

DEFAULT VALUE · 1

Colour

Sets the colour of the axis.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · Equally spaced colours from the `hcl` spectrum.

Ticks

Lets the user change graphical parameters associated with the ticks of the axis.

DETAIL · The ticks, together with the markers, calibrate the axis. The ticks are the short lines orthogonal to the axis, alongside which the markers (the numbers) are given. For non-linear axes, the slopes of the ticks are approximated numerically.

Desired number

Sets the desired number of ticks on the axis.

DETAIL · Corresponds to the argument `n` of R's `pretty` function. This number is not strictly adhered to; it is only approximate.

DEFAULT VALUE · 5

REFERENCES · See the R help file of the `pretty` function.

Line type

Sets the type of line used to draw the ticks on the axis.

DETAIL · Corresponds to R's `lty` argument. Possible values: 0, . . . , 6.

DEFAULT VALUE · 1

Line width

Sets the width of the ticks of the axis.

DETAIL · Corresponds to R's `lwd` argument. Larger values result in wider ticks. Possible values: positive real numbers.

DEFAULT VALUE · 1

Colour

Sets the colour of the ticks of the axis.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · The same as the default colour of the axis.

Relative length

Sets the relative length of the ticks of the axis.

DETAIL · The length of the ticks is twice this fraction of the width of the biplot.

DEFAULT VALUE · 0.0075

Markers

Lets the user change graphical parameters associated with the markers of the axis.

DETAIL · The markers, together with the ticks, calibrate the axis. The markers are the numbers, in terms of the original variable values, given alongside the ticks (the short lines orthogonal to the axis). The markers are positioned on that side of the axis that would place them below the axis if the axis were to be rotated to be horizontal (so that the marker values increase from left to right).

Font

Sets the font of the markers of the axis.

DETAIL · Corresponds to R's `font` argument. Possible values: 1: regular; 2: bold; 3: italic; 4: bold italic.

DEFAULT VALUE · 1

Size

Sets the size of the markers of the axis.

DETAIL · Corresponds to R's `cex` argument. Larger values result in larger text. Possible values: positive real numbers.

DEFAULT VALUE · 0.75

Colour

Sets the colour of the markers of the axis.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · The same as the default colour of the axis.

Relative offset

Sets the relative offset of the markers from the corresponding ticks of the axis.

DETAIL · The markers are shifted this fraction of the width of the biplot from the edge of the corresponding ticks.

DEFAULT VALUE · 0.005.

Label

Lets the user change the graphical parameters of the axis label.

Font

Sets the font of the axis label.

DETAIL · Corresponds to R's `font` argument. Possible values: 1: regular; 2: bold; 3: italic; 4: bold italic.

DEFAULT VALUE · 1

Size

Sets the size of the axis label.

DETAIL · Corresponds to R's `cex` argument. Larger values result in larger text. Possible values: positive real numbers.

DEFAULT VALUE · 0.75

Orientation

Sets the orientation of clinging axis labels.

DETAIL · Corresponds to R's `las` argument. Possible values: 0: always parallel to the biplot edge; 1: always horizontal; 2: always perpendicular to the biplot edge; 3: always vertical.

DEFAULT VALUE · 1

Colour

Sets the colour of the axis label.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · The same as the default colour of the axis.

Defaults

Reverts the options to their default values.

OK

Saves the options, returns to the GUI, redraws the currently displayed biplot region and other graphs if necessary.

Cancel

Cancels the changes made since the dialogue box was opened, and returns to the GUI.

▷ **Interaction...**

Lets the user set graphical parameters for variable value prediction and highlighted axes.

SCREENSHOT · 3.5

Prediction

Lets the user change the graphical parameters associated with variable value prediction.

Line type

Sets the line type used for projection.

DETAIL · Corresponds to R's `lty` argument. Possible values: 0, . . . , 6.

DEFAULT VALUE · 3

Line width

Sets the width of the projection lines.

DETAIL · Corresponds to R's `lwd` argument. Larger values result in wider ticks.
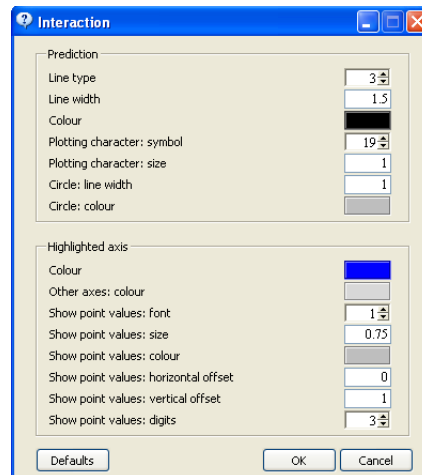
Screenshot 3.5: The *Format → Interaction* dialogue box.

Possible values: positive real numbers.

DEFAULT VALUE · 1.5
Colour
Sets the colour of the projection lines.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "black"
Plotting character: symbol
Sets the symbol used to represent the projected points on the axes.

DETAIL · Corresponds to R's `pch` argument. Possible values: `NA`, 0, . . . , 25.

DEFAULT VALUE · 19
Plotting character: size
Sets the size of the projected points on the axes.

DETAIL · Corresponds to R's `cex` argument. Larger values result in larger points. Possible values: positive real numbers.

DEFAULT VALUE · 1
Circle: line width
Sets the width of the circle used in circular projection in the circular non-linear biplot.

DETAIL · Corresponds to R's `lwd` argument. Larger values result in wider ticks. Possible values: positive real numbers.

DEFAULT VALUE · 1
Circle: colour
Sets the colour of the cirlce used in circular projection in the circular non-linear biplot.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "gray75"
Highlighted axis
Lets the user change the graphical parameters associated with highlighted

axes.

DETAIL · The graphical parameters prefixed with 'Show point values' are for the customisation of the variable values which are shown when an axis is highlighted and View → Show point values is selected.

Colour

Sets the colour of the highlighted axis.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "blue"

Other axes: colour

Sets the colour of the non-highlighted axes.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "gray85"

Show point values: font

Sets the font of the point values.

DETAIL · Corresponds to R's `font` argument. Possible values: 1: regular; 2: bold; 3: italic; 4: bold italic.

DEFAULT VALUE · 1

Show point values: size

Sets the size of the point values.

DETAIL · Corresponds to R's `cex` argument. Larger values result in larger text. Possible values: positive real numbers.

DEFAULT VALUE · 0.75

Show point values: colour

Sets the colour of the point values.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "gray75"

Show point values: horizontal offset

Sets the horizontal offset of the point values from the corresponding points.

DETAIL · The point values are shifted horizontally from the corresponding points with this multiple of the width of the letter 'x' (in the selected font size). Positive values shift the values to the right; negative values shift the values to the left.

DEFAULT VALUE · 0

Show point values: vertical offset

Sets the vertical offset of the point values from the corresponding points.

DETAIL · The point values are shifted vertically from the corresponding points with this multiple of the height of the letter 'x' (in the selected font size). Positive values shift the values towards the top; negative values shift the values towards the bottom.

DEFAULT VALUE · 1

Show point values: digits

Sets the number of decimal places to show in the point values.

DETAIL · Possible values: 0, . . . , 8.

DEFAULT VALUE · 3

Screenshot 3.6: The *Format* → *Diagnostics* dialogue box.

Defaults
Reverts the options to their default values.
OK
Saves the options, returns to the GUI, and redraws the currently displayed biplot region.
Cancel
Cancels the changes made since the dialogue box was opened, and returns to the GUI.

▷ **Diagnostic tabs...**
Lets the user set graphical parameters for the graphs of the diagnostic tabs.
SCREENSHOT · 3.6

Convergence
Lets the user change the graphical parameters of the graph of the convergence tab.
Line type
Sets the line type.
DETAIL · Corresponds to R's `lty` argument. Possible values: 0, ..., 6.
DEFAULT VALUE · 1
Line width
Sets the line width.
DETAIL · Corresponds to R's `lwd` argument. Possible values: positive real numbers.
DEFAULT VALUE · 1

Colour
Sets the line colour.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "red"

Predictivities
Lets the user change the graphical parameters of the points, groups and axes predictivities graphs of the points, group and axes tabs, respectively.

Plotting character: axes symbol
Sets the symbol used to represent axes in the axes tab.

DETAIL · Corresponds to R's `pch` argument. Possible values: `NA`, 0, . . . , 25.

DEFAULT VALUE · 19

Plotting character: size
Sets the size of the plotting characters.

DETAIL · Corresponds to R's `cex` argument. Larger values result in larger plotting characters. Possible values: positive real numbers.

DEFAULT VALUE · 1

Label: font
Sets the font of the labels.

DETAIL · Corresponds to R's `font` argument. Possible values: 1: regular; 2: bold; 3: italic; 4: bold italic.

DEFAULT VALUE · 1

Label: size
Sets the size of the labels.

DETAIL · Corresponds to R's `cex` argument. Larger values result in larger text. Possible values: positive real numbers.

DEFAULT VALUE · 0.85

Label: horizontal offset
Sets the horizontal offset of the labels from the corresponding plotting characters.

DETAIL · The labels are shifted horizontally from the corresponding plotting characters with this multiple of the width of the letter 'x' (in the selected font size). Positive values shift the labels to the right; negative values shift the labels to the left.

DEFAULT VALUE · 0

Label: vertical offset
Sets the vertical offset of the labels from the corresponding plotting characters.

DETAIL · The labels are shifted vertically from the corresponding plotting characters with this multiple of the height of the letter 'x' (in the selected font size). Positive values shift the labels towards the top; negative values shift the labels towards the bottom.
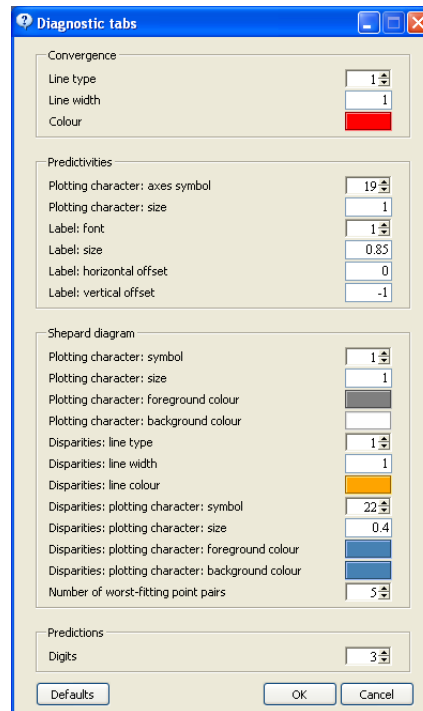
DEFAULT VALUE · −1

Shepard diagram
Lets the user change the graphical parameters of the Shepard diagram of the points tab.

Plotting character: symbol
Sets the symbol used to represent the inter-point distances.

DETAIL · Corresponds to R's `pch` argument. Possible values: `NA`, 0, ..., 25.

DEFAULT VALUE · 1

Plotting character: size
Sets the size of the plotting character used for inter-point distances.

DETAIL · Corresponds to R's `cex` argument. Larger values result in larger plotting characters. Possible values: positive real numbers.

DEFAULT VALUE · 1

Plotting character: foreground colour
Sets the exterior colour of the plotting character used for inter-point distances.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "gray50"

Plotting character: background colour
Sets the interior colour of the plotting character used for inter-point distances.

DETAIL · Corresponds to R's `bg` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "white"

Disparities: line type
Sets the line type used for inter-sample disparities.

DETAIL · Corresponds to R's `lty` argument. Possible values: 0, ..., 6.

DEFAULT VALUE · 1

Disparities: line width
Sets the width of the line used for inter-sample disparities.

DETAIL · Corresponds to R's `lwd` argument. Larger values result in a wider line. Possible values: positive real numbers.

DEFAULT VALUE · 1

Disparities: line colour
Sets the colour of the line used for inter-sample disparities.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "orange"

Disparities: plotting character: symbol
Sets the symbol used to represent the inter-sample disparities.

DETAIL · Corresponds to R's `pch` argument. Possible values: `NA`, 0, ..., 25.

DEFAULT VALUE · 22

Disparities: plotting character: size
Sets the size of the plotting character used for inter-samples disparities.

DETAIL · Corresponds to R's `cex` argument. Larger values result in larger plotting characters. Possible values: positive real numbers.

DEFAULT VALUE · 0.4

Disparities: plotting character: foreground colour
Sets the exterior colour of the plotting character used for inter-samples disparities.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the

colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "steelblue"

Disparities: plotting character: background colour

Sets the interior colour of the plotting character used for inter-samples disparities.

DETAIL · Corresponds to R's `bg` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "steelblue"

Number of worst-fitting point pairs

Sets the number of worst-fitting point pairs to show at the top left of the Shepard diagram.

DETAIL · The degree of misfit between a pair of points is quantified as the difference between the inter-point distance and the corresponding inter-sample disparity. Therefore as the vertical distance from the plotting character that represents the inter-point distance to the line representing the inter-sample disparities.

DEFAULT VALUE · 5

Predictions

Lets the user change the settings associated with the predictions tab.

Digits

Sets the number of decimal places to show in the variable value predictions of the predictions tab.

DETAIL · Possible values: 0, . . . , 8.

DEFAULT VALUE · 3

Defaults

Reverts the options to their default values.

OK

Saves the options, returns to the GUI, and redraws the the graphs of the diagnostic tabs.

Cancel

Cancels the changes made since the dialogue box was opened, and returns to the GUI.

▷ **Reset all. . .**

Lets the user revert all the parameters of the Format menu to their default values.

SHORTCUT · Ctrl+R

A standard dialogue box appears asking the user to confirm whether or not to reset the graphical parameters of the Format menu.

DETAIL · Upon confirmation, the graphical parameters are reset and the bipot region and the graphs of the diagnostic tabs are redrawn.

• **Joint**

For the construction of joint-mechanism biplots.

DETAIL · The biplots of this menu have both their points and axes determined according to a single, joint mechanism. Other biplots have their points determined from the Points menu and their axes determined from the Axes menu.

▷ **PCA**
Constructs a PCA biplot.

DETAIL · The points correspond to the scores of the first two PCs.

SHORTCUT · 1

ALTERNATIVE TO · Covariance/Correlation, CVA, None, Regression, Procrustes, Circular non-linear

REFERENCES · Gower and Hand (1996, Chapter 2)

▷ **Covariance/Correlation**
Contructs a covariance/correlation biplot.

DETAIL · If the data are only centred, a covariance biplot is produced. If the data are centred and scaled, a correlation biplot is produced.

SHORTCUT · 2

ALTERNATIVE TO · PCA, CVA, None, Regression, Procrustes, Circular non-linear

REFERENCES · Gabriel (1971), Greenacre (1984), Gardner (2001, Section 2.3.2), Gower and Hand (1996, Section 11.5.1)

▷ **CVA**
Constructs a CVA biplot.

DETAIL · The points correspond to the first two CVs. The group means can be included by clicking Additional → Interpolate → Sample group means. CVA biplots are invariant to the scaling of data to have unit variances. Only available when there is more than one group.

SHORTCUT · 3

ALTERNATIVE TO · PCA, Covariance/Correlation, None, Regression, Procrustes, Circular non-linear

REFERENCES · Gower and Hand (1996, Chapter 5 except Section 5.5)

• **Points**
Determines the positions of the points of those biplots which have their axes determined from the Axes menu.

DETAIL · As opposed to the biplots of the Joint menu, the Points and Axes menus are used together. The options of the Points menu determine the points, while the options of the Axes menu determine the axes.

▷ **Distance metric**
Lets the user set the distance metric to be used in PCO and MDS.

DETAIL · All the distance metrics that are available are Euclidean-embeddable. Therefore points exist in Euclidean space which have the same distances between them in Pythagorean distance as the samples have between them in the Euclidean-embeddable distance metric. If the axes previously selected in the Axes menu are no longer available because of the change in distance metric, the most appropriate replacement is automatically chosen.

REFERENCES · Gower and Legendre (1986), Gower and Hand (1996, Section A.5.1)

○ **Pythagoras**
The distance metric is set to Pythagoras distance.

DETAIL · 'Pythagoras distance' and 'Euclidean distance' are synomomous. However, to avoid confusion with the class of 'Euclidean-embeddable' distances, the terminology 'Pythagoras distance' is used instead. Under this distance

metric, the distance between a vector $\mathbf{x} : p \times 1$ and a vector $\mathbf{y} : p \times 1$ is given by $((\mathbf{x} - \mathbf{y})'(\mathbf{x} - \mathbf{y}))^{1/2}$.

ALTERNATIVE TO · Square-root-of-Manhattan, Clark, Mahalanobis

REFERENCES · Loomis (1968), Cox and Cox (2001, p. 11), Borg and Groenen (2005, p. 122)

○ **Square-root-of-Manhattan**

The distance metric is set to Square-root-of-Manhattan distance.

DETAIL · Distances under this metric are the square root of those under the Manhattan or City Block distance metric. Therefore the distance between a vector $\mathbf{x} : p \times 1$ and a vector $\mathbf{y} : p \times 1$ is given by $(\sum_{i=1}^{p} |x_i - y_i|)^{1/2}$.

ALTERNATIVE TO · Pythagoras, Clark, Mahalanobis

REFERENCES · Gower and Harding (1988), Gower and Hand (1996, p. 103), Cox and Cox (2001, p. 11), Borg and Groenen (2005, p. 122)

○ **Clark**

The distance metric is set to Clark distance.

DETAIL · Under this distance metric, the distance between a vector $\mathbf{x} : p \times 1$ and a vector $\mathbf{y} : p \times 1$ is given by $\sum_{i=1}^{p} \left(\frac{x_i - y_i}{x_i + y_i} \right)^2$.

ALTERNATIVE TO · Pythagoras, Square-root-of-Manhattan, Mahalanobis

REFERENCES · Clark (1952), Gower and Ngouenet (2005)

○ **Mahalanobis**

The distance metric is set to Mahalanobis distance.

DETAIL · Under this distance metric, the distance between a vector $\mathbf{x} : p \times 1$ and a vector $\mathbf{y} : p \times 1$ is given by $((\mathbf{x} - \mathbf{y})'\mathbf{S}^{-1}(\mathbf{x} - \mathbf{y}))^{1/2}$ where $\mathbf{S}$ is the covariance matrix of the matrix with all the observations as rows. Mahalanobis distance is also known as 'statistical distance'.

ALTERNATIVE TO · Pythagoras, Square-root-of-Manhattan, Clark

REFERENCES · Mahalanobis (1936), Cox and Cox (2001, p. 11)

▷ **PCO**

The points are determined by PCO, with inter-sample dissimilarities calculated according to the chosen distance metric.

DETAIL · The coordinates of the points are taken to be the first two principal coordinates. PCO is also known as 'classical scaling'.

SHORTCUT · A

ALTERNATIVE TO · MDS Identity transformation, MDS Monotone regression, MDS Monotone spline transformation.

REFERENCES · Gower (1966), Gower and Hand (1996, Section A.5.2), Cox and Cox (2001, Section 2.2), Borg and Groenen (2005, Chapter 12)

▷ **MDS**

The points are determined by MDS, with inter-sample dissimilarities calculated according to the chosen distance metric.

REFERENCES · Cox and Cox (2001), Borg and Groenen (2005)

○ **Random initial configuration**

The next MDS algorithm will start from a random initial configuration.

DETAIL · If selected, the initial MDS configuration is taken to have points uniformly distributed over $[-1, 1]$ in both dimensions. If not selected, the initial configuration is taken to be the previous configuration of points. Useful to avoid local minima.
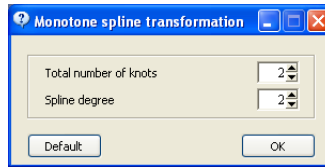
Screenshot 3.7: The *Points → MDS → Monotone spline transformation* dialogue box.

○ **Identity transformation**

Performs an MDS with the identity transformation.

DETAIL · The inter-sample disparities are taken to be the inter-sample dissimilarities, calculated according to the chosen distance metric. Using an IM algorithm, the points are chosen so that stress, the sum of squared differences between the inter-point distances and the inter-sample disparities, is minimised.

SHORTCUT · B

ALTERNATIVE TO · PCO, Monotone regression, Monotone spline transformation

REFERENCES · Borg and Groenen (2005, Section 8.6)

○ **Monotone regression**

Performs a non-metric MDS.

DETAIL · The inter-sample disparities are calculated by the up-and-down-blocks algorithm from the inter-sample dissimilarities, based on the chosen distance metric. In essense, merely the order of the inter-sample dissimilarities is preserved. Using an IM algorithm, the points are chosen so that stress, the sum of squared differences between the inter-point distances and the inter-sample disparities, is minimised. The Primary approach to ties or the Secondary approach to ties can be used.

SHORTCUT · C

ALTERNATIVE TO · PCO, Identity transformation, Monotone spline transformation,

REFERENCES · Borg and Groenen (2005, Sections 9.1, 9.2)

○ **Monotone spline transformation. . .**

Performs an MDS with a monotone spline transformation

DETAIL · The inter-sample dissimilarities are smoothed using monotone splines to obtain the inter-sample disparities, based on the chosen distance metric. Using an IM algorithm, the points are chosen so that stress, the sum of squared differences between the inter-point distances and the inter-sample disparities, is minimised.

SHORTCUT · D

SCREENSHOT · 3.7

ALTERNATIVE TO · PCO, Identity transformation, Monotone regression

REFERENCES · Borg and Groenen (2005, Sections 9.1, 9.6)

Lets the user select the spline transformation.
Total number of knots
The total number of spline knots, including the two exterior knots.

DETAIL · Between each successive pair of knots, a polynomial of degree

> *Spline degree* is fit.
>
> DEFAULT VALUE · 2
>
> REFERENCES · Borg and Groenen (2005, Section 9.6)
>
> **Spline degree**
>
> The degree of the polynomial of the spline transformation.
>
> DETAIL · Between each successive pair of knots, a polynomial of degree *Spline degree* is fit.
>
> DEFAULT VALUE · 2
>
> REFERENCES · Borg and Groenen (2005, Section 9.6)
>
> **Default**
>
> Reverts the options to their default values.
>
> **OK**
>
> Saves the options, returns to the GUI, and starts the MDS iteration process.
>
> **Cancel**
>
> Cancels the changes made since the dialogue box was opened, and returns to the GUI.

- ○ **Primary approach to ties**
  If selected, equality in dissimilarities need not translate to equality in disparities in monotone regression.

  DETAIL · Only available when Monotone regression is selected.

  ALTERNATIVE TO · Secondary approach to ties

  REFERENCES · Borg and Groenen (2005, Section 9.4)

- ○ **Secondary approach to ties**
  If selected, equality in dissimilarities implies equality in disparities in monotone regression.

  DETAIL · Only available when Monotone regression is selected.

  ALTERNATIVE TO · Primary approach to ties

  REFERENCES · Borg and Groenen (2005, Section 9.4)

- ○ **In terms of principal axes**
  If selected, newly converged MDS configurations are rotated to be in terms of their principal axes.

  DETAIL · Since scaling techniques are invariant to rotation (amongst other things), expressing MDS solutions in terms of their principal axes can make different configurations more comparable.

  REFERENCES · Borg and Groenen (2005, Section 24.1)

- • **Axes**
  Determines the axes to be shown along with the points determined from the Points menu.

  DETAIL · As opposed to the biplots of the Joint menu, the Points and Axes menus are used together. The options of the Points menu determine the points, while the options of the Axes menu determine the axes. The available axes depend on the choice of distance metric and the scaling technique from the Points menu.

  - ▷ **None**
    No biplot axes are shown, only points.

    SHORTCUT · 0

    ALTERNATIVE TO · PCA, Covariance/Correlation, CVA, Regression, Procrustes, Circular non-linear

▷ **Regression**
Shows regression biplot axes.

DETAIL · Not available when the distance metric is Mahalanobis and the scaling technique is PCO.

SHORTCUT · 4

ALTERNATIVE TO · PCA, Covariance/Correlation, CVA, None, Procrustes, Circular non-linear

REFERENCES · Gower and Hand (1996, Sections 3.3.2, 3.4.3), Cox and Cox (2001, Section 3.7)

▷ **Procrustes**
Shows Procrustes biplot axes.

DETAIL · Not available when the distance metric is Mahalanobis and the scaling technique is PCO.

SHORTCUT · 5

ALTERNATIVE TO · PCA, Covariance/Correlation, CVA, None, Regression, Circular non-linear

REFERENCES · Gower and Hand (1996, Sections 3.3.1, 3.4.2, A.10.2), Gower and Dijksterhuis (2004, p. 57)

▷ **Circular non-linear**
Shows circular non-linear biplot axes.

DETAIL · Only available when the points are determined by PCO, except when the distance metric is Mahalanobis.

SHORTCUT · 6

ALTERNATIVE TO · PCA, Covariance/Correlation, CVA, None, Regression, Procrustes

REFERENCES · Gower and Harding (1988), Gower and Hand (1996, Chapter 6), Gower and Ngouenet (2005)

▷ **Default**
Selects the default biplot type from the Axes menu.

DETAIL · The default biplot type depends on the choice of distance metric and scaling technique from the Points menu.

• **Additional**
Contains additional descriptors which can be included in the biplot.

DETAIL · When the entries of the Additional menu appear in the legend, the group names are preceded by short prefixes. The prefixes are as follows: SGM for 'Sample Group Mean'; CH for 'Convex Hull'; AB for 'Alpha-Bag'; TM for 'Tukey Median'; and CR for 'Classification Region'.

▷ **Interpolate**
Lets the user interpolate additional points onto the biplot.

DETAIL · Only a single new point can be interpolated at one time.

○ **A new sample. . .**
Lets the user interpolate a new sample as a point in the biplot.

DETAIL · Not available if there are no axes, as opposed to the axes being hidden.

SHORTCUT · Ctrl+N

SCREENSHOT · 3.8

REFERENCES · Gower and Hand (1996, Sections 2.3, 6.3.1)
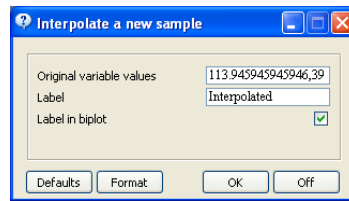
Screenshot 3.8: The *Additional → Interpolate → A new sample* dialogue box.

Original variable values
The variable values of the sample which is to be interpolated.

DETAIL · The variable values are given in terms of the units of the original variables, and in the order of the original variables. The variable values are separated by commas, without spaces. The variable values of variables that have been dragged from the biplot into the kraal are not entered. If axes are dragged back onto the biplot from the kraal, the option is disabled.

DEFAULT VALUE · The mean values of the non-kraal variables.

Label
The label to be used to annotate the interpolated point in the biplot or in the legend.

DEFAULT VALUE · Interpolated

Label in biplot
If checked, shows the interpolated point's label in the biplot itself. Otherwise the label is shown in the legend.

DEFAULT VALUE · checked

Defaults
Reverts the options to their default values.

Format
See the next framed section.

OK
Saves the options, returns to the GUI, and redraws the biplot region with the interpolated point.

Off
Disables the option, and returns to the GUI, not showing the interpolated point.

Format
Lets the user set graphical parameters for interpolated samples.

SCREENSHOT · 3.9

Plotting character
Lets the user change the graphical parameters of the of the interpolated point.

Symbol
Sets the symbol used to represent the interpolated point.

DETAIL · Corresponds to R's `pch` argument. Possible values: `NA`, 0, . . . , 25.

DEFAULT VALUE · 22
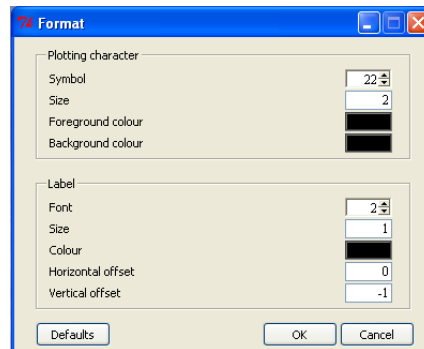
Screenshot 3.9: The *Additional → Interpolate → A new sample → Format* dialogue box.

> **Size**
> Sets the size of the interpolated point.
>
> DETAIL · Corresponds to R's `cex` argument. A larger value results in a larger point. Possible values: positive real numbers.
>
> DEFAULT VALUE · 2
> **Foreground colour**
> Sets the exterior colour of the interpolated point.
>
> DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.
>
> DEFAULT VALUE · "black"
> **Background colour**
> Sets the interior colour of the interpolated point.
>
> DETAIL · Corresponds to R's `bg` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.
>
> DEFAULT VALUE · "black"
> **Label**
> Lets the user change the graphical parameters of the interpolated point label.
> **Font**
> Sets the font of the interpolated point label.
>
> DETAIL · Corresponds to R's font argument. Possible values: 1: regular; 2: bold; 3: italic; 4: bold italic.
>
> DEFAULT VALUE · 2
> **Size**
> Sets the size of the interpolated point label.
>
> DETAIL · Corresponds to R's `cex` argument. Larger values result in larger text. Possible values: positive real numbers.
>
> DEFAULT VALUE · 1
> **Colour**
> Sets the colour of the interpolated point label.
>
> DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.
>
> DEFAULT VALUE · "black"
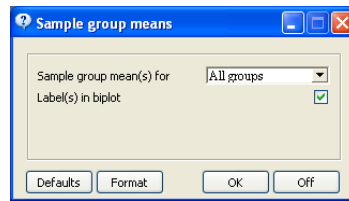
Screenshot 3.10: The *Additional → Interpolate → Sample group means* dialogue box.

> **Horizontal offset**
> Sets the horizontal offset of the interpolated point label from the interpolated point.
>
> DETAIL · The interpolated point label is shifted horizontally from the interpolated point with this multiple of the width of the letter 'x' (in the selected font size). Positive values shift the label to the right; negative values shift the label to the left.
>
> DEFAULT VALUE · 0
> **Vertical offset**
> Sets the vertical offset of the interpolated point label from the interpolated point.
>
> DETAIL · The interpolated point label is shifted vertically from the interpolated point with this multiple of the height of the letter 'x' (in the selected font size). Positive values shift the label towards the top; negative values shift the label towards the bottom.
>
> DEFAULT VALUE · −1
> **Defaults**
> Reverts the options to their default values.
> **OK**
> Saves the options and returns to the Additional → Interpolate → A new sample dialogue box.
> **Cancel**
> Cancels the changes made since the dialogue box was opened, and returns to the Additional → Interpolate → A new sample dialogue box.

○ **Sample group means. . .**
Lets the user interpolate the sample group means as points in the biplot.

DETAIL · This option is especially important in the case of CVA biplots, where the optimality criterion is in terms of separation of the group means. If the sample group mean labels are shown in the legend, the group names are preceded by the prefix SGM for 'Sample Group Mean'. Not available if there are no axes, as opposed to the axes being hidden.

SCREENSHOT · 3.10

REFERENCES · Gower and Hand (1996, Sections 2.3, 6.3.1)

> **Sample group mean(s) for**
> Sets the group(s) of samples for which sample group means must be interpolated.
>
> DETAIL · Possible options: All samples (the overall mean is interpolated),
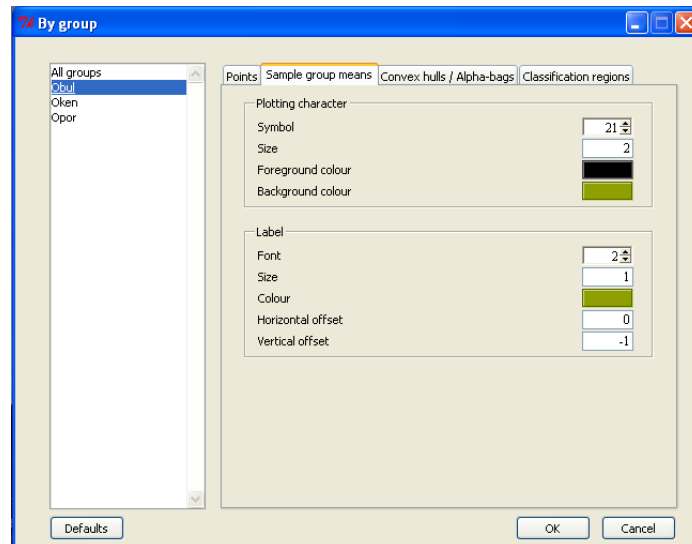
Screenshot 3.11: The *Additional → Interpolate → Sample group means → Format* dialogue box. This is also the *Sample group means* tab of the *Format → By group* dialogue box

all groups (the mean of each group is interpolated), or a specific group.

DEFAULT VALUE · All groups

Label(s) in biplot

If checked, shows the sample group mean labels in the biplot itself. Otherwise the labels are shown in the legend.

DEFAULT VALUE · checked

Defaults

Reverts the options to their default values.

Format

See the next framed region.

OK

Saves the options, returns to the GUI, and redraws the currently displayed biplot region with the interpolated sample group mean(s).

Off

Disables the option, and returns to the GUI, not showing sample group mean(s).

Format

Lets the user set graphical parameters for interpolated sample group means, for a particular group, or for all groups simultaneously.

SCREENSHOT · 3.11

This is also the Sample group means tab of the Format → By group dialogue box.

Plotting character

Lets the user change the graphical parameters of the point that represents the sample group mean, by group.

Symbol

Sets the symbol used to represent the sample group mean, by group.

DETAIL · Corresponds to R's `pch` argument. Possible values: `NA`, 0, ..., 25.

DEFAULT VALUE · If there is only one group, 22. Otherwise the values are recycled by group with the values 21 to 25.

Size

Sets the size of the point that represents the sample group mean, by group.

DETAIL · Corresponds to R's `cex` argument. A larger value results in a larger point. Possible values: positive real numbers.

DEFAULT VALUE · 2

Foreground colour

Sets the exterior colour of the point that represents the sample group mean, by group.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · "black"

Background colour

Sets the interior colour of the sample group mean, by group.

DETAIL · Corresponds to R's `bg` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · If there is only one group, "red". Otherwise equally spaced colours from the `hcl` spectrum.

Label

Lets the user change the graphical parameters of the sample group mean label, by group.

Font

Sets the font of the sample group mean label, by group.

DETAIL · Corresponds to R's `font` argument. Possible values: 1: regular; 2: bold; 3: italic; 4: bold italic.

DEFAULT VALUE · 2

Size

Sets the size of the sample group mean label, by group.

DETAIL · Corresponds to R's `cex` argument. Larger values result in larger text. Possible values: positive real numbers.

DEFAULT VALUE · 1

Colour

Sets the colour of the sample group mean label, by group.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · If there is only one group, "black". Otherwise equally spaced colours from the `hcl` spectrum.

Horizontal offset

Sets the horizontal offset of the sample group mean label from the corresponding sample group mean, by group.

DETAIL · The sample group mean label is shifted horizontally from the sample group mean with this multiple of the width of the letter 'x' (in the

Screenshot 3.12: The *Additional → Convex hulls* dialogue box.

---

selected font size). Positive values shift the labels to the right; negative values shift the labels to the left.

DEFAULT VALUE · 0

Vertical offset

Sets the vertical offset of the sample group mean label from the corresponding sample group mean, by group.

DETAIL · The sample group mean label is shifted vertically from the sample group mean with this multiple of the height of the letter 'x' (in the selected font size). Positive values shift the label towards the top; negative values shift the label towards the bottom.

DEFAULT VALUE · −1

Defaults

Reverts the options of all the tabs of the dialogue box to their default values.

OK

Saves the options of all the tabs of the dialogue box, and returns to the Additional → Interpolate → Sample group means dialogue box.

Cancel

Cancels the changes made since the dialogue box was opened, and returns to the Additional → Interpolate → Sample group means dialogue box.

▷ **Convex hulls. . .**

Lets the user superimpose convex hulls around one or more of the groups of points.

DETAIL · Convex hulls may be drawn around all the points, around each group of points, or around a specific group of points. If the convex hull labels are shown in the legend, the group names are preceded by the prefix CH for 'Convex Hull'.

SCREENSHOT · 3.12

ALTERNATIVE TO · Alpha-bags

REFERENCES · See the R help file of the `chull` function.

Convex hull(s) for

The group(s) of points for which convex hull(s) must be drawn.

DETAIL · Possible options: All points (a single convex hull is drawn around all points), All groups (a convex hull is drawn around each group), or a specific group.

DEFAULT VALUE · All groups

Defaults

Reverts the options to their default values.
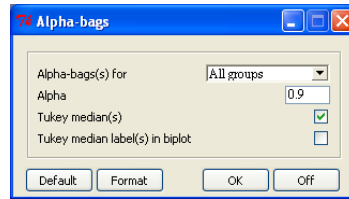
Screenshot 3.13: The *Additional* → *Alpha-bags* dialogue box.

Format
Discussed under Alpha-bags.
OK
Saves the option, returns to the GUI, and redraws the biplot region with the convex hull(s).
Off
Disables the option, and returns to the GUI, not showing convex hulls.

▷ **Alpha-bags...**
Lets the user superimpose alpha-bags around one or more of the groups of points.
DETAIL · Alpha-bags may be drawn around all the points, around each group of points, or around a specific group of points. Alpha-bags are closely related to bagplots which can be considered to be two-dimensional boxplots. Alpa-bags enclose approximately the inner $100\alpha\%$ of the points of a group. If the alpha-bag labels are shown in the legend, the group names are preceded by the prefix AB for 'Alpha-Bag'. If the Tukey median labels are shown in the legend, the group names are preceded by the prefix TM for 'Tukey Median'.
SCREENSHOT · 3.13
ALTERNATIVE TO · Convex hulls
REFERENCES · Rousseeuw *et al.* (1999), Gardner (2001), Wurz *et al.* (2003)

Alpha-bag(s) for
The group(s) of points for which alpha-bag(s) must be drawn.
DETAIL · Possible options: All points (a single alpha-bag is drawn for all points), All groups (an alpha-bag is drawn for each group), or a specific group.
DEFAULT VALUE · All groups
Alpha
The value of alpha for the alpha-bag(s) to be shown.
DETAIL · Approximately the inner $100\alpha\%$ of the points of the group are enclosed in an alpha-bag.
DEFAULT VALUE · 0.9
Tukey median(s)
If checked, the tukey median(s) of the groups are also shown.
DEFAULT VALUE · checked
REFERENCES · Rousseeuw *et al.* (1999), Gardner (2001)
Tukey median label(s) in biplot
If checked and Tukey median(s) are shown, the Tukey median labels are shown in the biplot, rather than in the legend.
DETAIL · Approximately the inner $100\alpha\%$ of the points of the group are enclosed
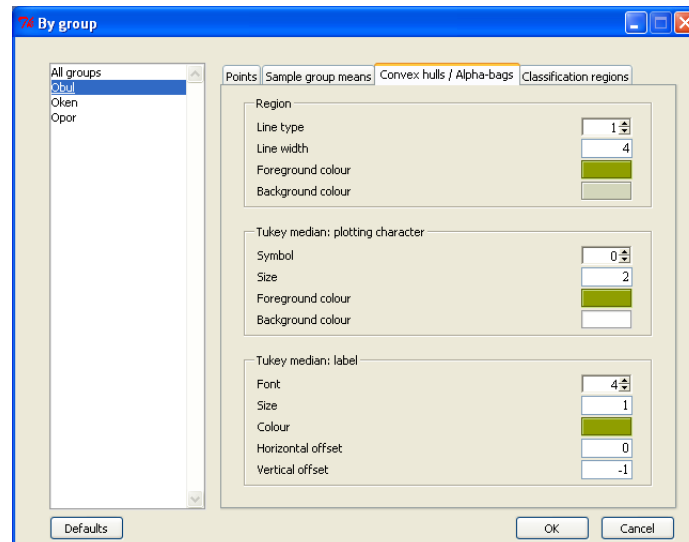
Screenshot 3.14: The *Additional → Convex hulls / Alpha-bags → Format* dialogue box. This is also the *Convex hulls / Alpha-bags* tab of the *Format → By group* dialogue box.

---

in an alpha-bag.

DEFAULT VALUE · clear

Defaults

Reverts the options to their default values.

Format

See the next framed section.

OK

Saves the options, returns to the GUI and redraws the biplot region with the alpha-bag(s) as selected.

Off

Disables the option, and returns to the GUI, not showing alpha-bags.

Format

Lets the user set graphical parameters for convex hulls and alpha-bags, for a particular group, or for all groups simultaneously.

SCREENSHOT · 3.14

This is also the Convex hulls / Alpha-bags tab of the Format → By group dialogue box.

Region

Lets the user change graphical parameters associated with convex hulls or alpha-bags.

Line type

Sets the type of line used to draw the convex hull or alpha-bag, by group.

DETAIL · Corresponds to R's `lty` argument. Possible values: 0, . . . , 6.

DEFAULT VALUE · 1

Line width

Sets the width of the convex hull or alpha-bag, by group.

DETAIL · Corresponds to R's `lwd` argument. Larger values result in wider

---

border for convex hulls and alpha-bags. Possible values: positive real numbers.

DEFAULT VALUE · 4

Foreground colour

Sets the exterior colour of the convex hull or alpha-bag, by group.

DETAIL · Corresponds to R's `col` argument.

DEFAULT VALUE · If there is only one group, #919191. Otherwise equally spaced colours from the `hcl` spectrum.

Background colour

Sets the interior colour of the convex hull or alpha-bag, by group.

DETAIL · Corresponds to R's `bg` argument.

DEFAULT VALUE · If there is only one group, #D4D4D4. Otherwise equally spaced colours from the `hcl` spectrum.

Tukey median: Plotting character

Lets the user change the graphical parameters of the Tukey median, by group.

Symbol

Sets the symbol used to represent the Tukey median, by group.

DETAIL · Corresponds to R's `pch` argument. Possible values: `NA`, 0, . . . , 25.

DEFAULT VALUE · 0

Size

Sets the size of the Tukey medians, by group.

DETAIL · Corresponds to R's `cex` argument. A larger value results in a larger point. Possible values: positive real numbers.

DEFAULT VALUE · 2

Foreground colour

Sets the exterior colour of the Tukey median, by group.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · If there is only one group, "red". Otherwise equally spaced colours from the `hcl` spectrum.

Background colour

Sets the interior colour of the Tukey median, by group.

DETAIL · Corresponds to R's `bg` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · `NA`

Tukey median: label

Lets the user change the graphical parameters of the Tukey median label, by group.

Font

Sets the font of the Tukey median label, by group.

DETAIL · Corresponds to R's `font` argument. Possible values: 1: regular; 2: bold; 3: italic; 4: bold italic.

DEFAULT VALUE · 4

Size

Sets the size of the Tukey median label, by group.

DETAIL · Corresponds to R's `cex` argument. Larger values result in larger text. Possible values: positive real numbers.

DEFAULT VALUE · 1
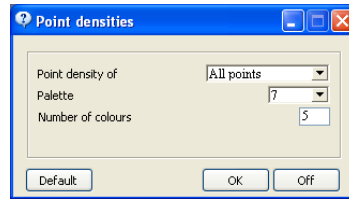
Screenshot 3.15: The *Additional → Point densities* dialogue box.

Colour

Sets the colour of the Tukey median label, by group.

DETAIL · Corresponds to R's `col` argument. Can be changed by clicking the colour, producing a standard colour selection dialogue box.

DEFAULT VALUE · If there is only one group, "black". Otherwise equally spaced colours from the `hcl` spectrum.

Horizontal offset

Sets the horizontal offset of the Tukey median label from the corresponding Tukey median, by group.

DETAIL · The Tukey median label is shifted horizontally from the Tukey median with this multiple of the width of the letter 'x' (in the selected font size). Positive values shift the labels to the right; negative values shift the labels to the left.

DEFAULT VALUE · 0

Vertical offset

Sets the vertical offset of the Tukey median label from the corresponding Tukey median.

DETAIL · The Tukey median label is shifted vertically from the Tukey median with this multiple of the height of the letter 'x' (in the selected font size). Positive values shift the label towards the top; negative values shift the label towards the bottom.

DEFAULT VALUE · −1

Defaults

Reverts the options of all the tabs of the dialogue box to their default values.

OK

Saves the options of all the tabs of the dialogue box, and returns to the Additional → Convex hulls or Additional → Alpha-bags dialogue box.

Cancel

Cancels the changes made since the dialogue box was opened, and returns to the Additional → Convex hulls or Additional → Alpha-bags dialogue box.

▷ **Point densities. . .**

Lets the user superimpose the biplot onto a background of point density estimates.

SCREENSHOT · 3.15

ALTERNATIVE TO · Classification regions

REFERENCES · Wand (2008), Blasius *et al.* (2008)
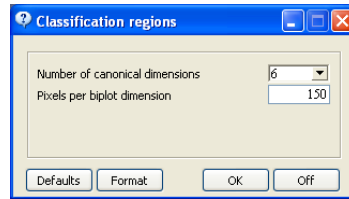
Screenshot 3.16: The *Additional → Classification regions* dialogue box.

> **Point density of**
> Sets the group of points on which the point densities are to be based.
>
> DETAIL · Possible options: All points, or a specific group of points.
>
> DEFAULT VALUE · All points
>
> **Palette**
> The colour palette in which the point densities are represented.
>
> DETAIL · Possible options: Terrain, Heat, 1, . . . , 8
>
> DEFAULT VALUE · 7
>
> **Number of colours**
> The number of distinct colours to include in the palette.
>
> DETAIL · The higher the number, the smoother the image. The lower the number, the easier to distinguish between different levels.
>
> DEFAULT VALUE · 5
>
> **Defaults**
> Reverts the options to their default values.
>
> **OK**
> Saves the options, returns to the GUI and redraws the biplot region with the point density estimates as chosen.
>
> **Off**
> Disables the option, and returns to the GUI, not showing point density estimates.

▷ **Classification regions. . .**
Lets the user superimpose a CVA biplot onto a background of coloured classification regions.

DETAIL · Pixels are coloured in according to the closest group mean in a specified number of canonical dimensions. If the classification region labels are shown in the legend, the group names are preceded by the prefix CR for 'Classification Region'. Only available for CVA biplots.

SCREENSHOT · 3.16

ALTERNATIVE TO · Point densities

REFERENCES · Gower and Hand (1996, Section 7.3.2), Gardner and Le Roux (2005)

> **Number of canonical dimensions**
> The number of canonical dimensions in which classification are to be made.
>
> DETAIL · From 1, . . . , min(the number of variables, one less than the number of groups).
>
> DEFAULT VALUE · The maximum value.
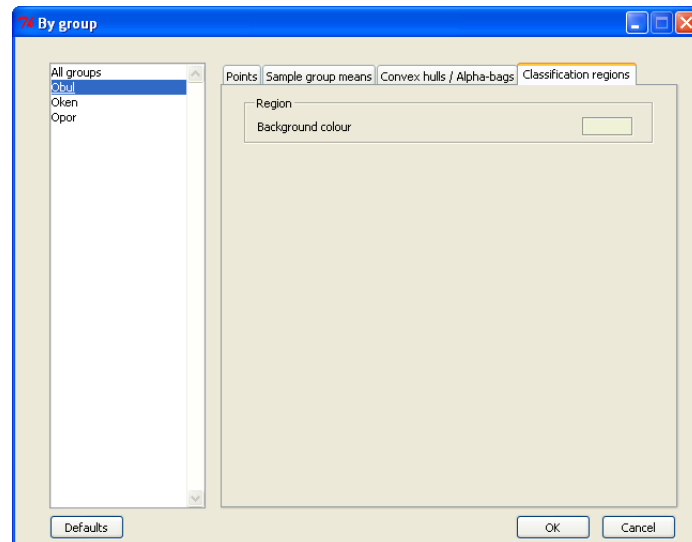
Screenshot 3.17: The *Additional → Classification regions → Format* dialogue box. This is also the *Classification regions* tab in the *Format → By group* dialogue box.

Pixels per biplot dimension
This number of pixels, squared, are classified into one of the groups and colour-coded accordingly.

DETAIL · The larger the number, the finer the image, but the more time and memory required to produce it. The smaller the number, the grainier the image.

DEFAULT VALUE · 150

Default
Reverts the option to its default value.

Format
See the next framed region.

OK
Saves the options, returns to the GUI and redraws the biplot region with the classification regions shown.

Off
Disables the option, and returns to the GUI, without classification regions.

Format
Lets the user set graphical parameters for classification regions, for a particular group, or for all groups simultaneously.

SCREENSHOT · 3.17

This is also the Classification regions tab in the Format → By group dialogue box.

Region
Background colour
The background colour of classification region, by group.

DETAIL · Equally spaced colours from the `hcl` spectrum.

> **Defaults**
> Reverts the options to their default values.
> **OK**
> Saves the options, returns to the GUI and redraws the biplot region with classification regions.
> **Off**
> Disables the option, and returns to the GUI, not showing classification regions.

> ▷ **Clear all**
> Disables all additional descriptors.
>
> SHORTCUT · Ctrl+L

- **Help**
  Some help options.

  > ▷ **Manual (in PDF)**
  > Opens this manual.
  >
  > SHORTCUT · F1

  > ▷ **Homepage**
  > Opens the package homepage at `http://biplotgui.r-forge.r-project.org/` if an Internet connection is available.

  > ▷ **Report a bug**
  > Opens the webpage `http://r-forge.r-project.org/tracker/?group˙id=225` if an Internet connection is available.

  > ▷ **Show pop-up help**
  > If selected, pop-up help is shown as the mouse cursor is hovered over the major elements of the main view of the GUI.

  > ▷ **About. . .**
  > Shows information about the author, version and license.

▶ **Biplot region**
Contains the biplot, optional title, and optional legend.

DETAIL · The biplot region is located towards the left of the GUI. The biplot itself is positioned at the centre of the biplot region. An optional title is shown above the biplot when View → Show title is selected. An optional legend may be shown below the biplot.

The biplot region is responsive to mouse clicks and motion. Different pop-up menus appear when right clicking inside the biplot, on a point, on an axis, or outside the biplot.

Depending on how the GUI has been resized onscreen, the proportions within the biplot region onscreen may differ from the proportions in biplot regions that have been saved, copied, printed or shown in external windows. To ensure consistency, saved and copied biplot regions always have both their width and height set to 20.32 centimeters (8 inches). Two-dimensional external-window biplot regions always have both their width and height set to 17.78 centimeters (7 inches). This is the default size of R graphs.

SCREENSHOT · 2.1

- **Title**
  Provides a title to the biplot.

  DETAIL · The title is only shown if View → Show title is selected. The default title depends on the currently displayed biplot. The biplot retains the same size and relative position, irrespective of whether or not a title is shown.

- **Biplot**

  A graph that displays samples as points, and variables as calibrated axes.

  DETAIL · This is the graph of main interest. Right clicking on a point, on an axis, or elsewhere within the biplot produces pop-up menus. Points and axes may be dragged from the biplot into the kraal, and be dragged back onto the biplot from the kraal.

  REFERENCES · Gabriel (1971), Gower and Hand (1996)

- **Legend**

  A legend to the elements which appear in the biplot.

  DETAIL · If necessary, a legend is provided below the biplot. The legend can identify groups, axes, an interpolated point, sample group means, convex hulls, alpha-bags, Tukey medians and classification regions, in that order. All these elements can separately be included or excluded from the legend.

  The legend can contain a maximum of 16 entries at any one time. If there are more than 16 entries, only the first 16 are shown initially. The succeeding set of entries can be shown by right clicking outside the biplot but inside the biplot region, and selecting Show next legend entries from the popup-menu. The keyboard shortcut is Ctrl++. Similar options are available with which to show the preceding entries. All legend entries are restricted to be 14 characters long.

  When the entries of the Additional menu appear in the legend, the group names are preceded by short prefixes. The prefixes are as follows: SGM for Sample Group Mean; CH for Convex Hull; AB for Alpha-Bag; TM for Tukey Median; CR for Classification Region.

- **Inner pop-up menu**

  Provides access to biplot options.

  DETAIL · This pop-up menu appears when the user right clicks inside the biplot (as opposed to the biplot region), but not on a point or on an axis.

  ▷ **Zoom in**

  Zooms in around the right clicked position.

  DETAIL · The biplot is enlarged and the focus is shifted so that that position which was right clicked to produce the pop-up menu appears at the centre of the biplot.

  ▷ **Zoom out**

  Zooms out from the right clicked position.

  DETAIL · The biplot is reduced in size and the focus is shifted so that that position which was right clicked to produce the pop-up menu appears at the centre of the biplot.

  ▷ **Reset zoom**

  Resets the biplot to its original zoom factor.

  ▷ **Don't predict**

  Dynamic variable prediction is disabled.

  DETAIL · Not shown if there are no axes, as opposed to the axes being hidden.

  ALTERNATIVE TO · Predict cursor positions, Predict points closest to cursor positions

  ▷ **Predict cursor positions**

  The variable values of the point under the cursor position are dynamically predicted.

  DETAIL · For linear biplots, the orthogonal projections onto the biplot axes are shown. For non-linear biplots, circular projections are shown. The predicted variable values themselves are shown in the predictions tab. Not available when

the axes are hidden, or if the biplot is interpolative rather than predictive. Not shown when there are no axes.

ALTERNATIVE TO · Don't predict, Predict points closest to cursor positions

REFERENCES · (Gower and Hand, 1996, Sections 2.3, 6.3.2)

▷ **Predict points closest to cursor positions**
The variable values of the point closest to the cursor position are dynamically predicted.

DETAIL · For linear biplots, orthogonal projections onto the biplot axes are shown. For non-linear biplots, circular projections are shown. The predicted variable values themselves are shown in the Predictions tab. Not available when the points are hidden, the axes are hidden, or if the biplot is interpolative rather than predictive. Not shown when there are no axes.

ALTERNATIVE TO · Don't predict, Predict cursor positions

REFERENCES · (Gower and Hand, 1996, Sections 2.3, 6.3.2)

▷ **Remove axis highlight**
If an axis has been highlighted, this option removes the highlight.

DETAIL · Only available if an axis has been highlighted. Axes are highlighted by right clicking them and selecting Highlight from the pop-up menu. Not shown when there are no axes, as opposed to the axes being hidden.

▷ **Format by group...**
The same as Format → By group.

▷ **Format axes...**
The same as Format → Axes.

DETAIL · Not shown when there are no axes, as opposed to the axes being hidden.

▷ **Save as**
The same as File → Save as.

▷ **Copy**
The same as File → Copy.

▷ **Print...**
The same as File → Print.

• **Point pop-up menu**
Provides access to options associated with points.

DETAIL · This pop-up menu appears when the user right clicks a point inside the biplot. A point or axis is under the mouse if the mouse cursor changes from an arrow to a hand.

▷ **Send to kraal**
Sends the point which was right clicked to the kraal.

DETAIL · Sending a point to the kraal temporarily removes it from consideration. The biplot region and diagnostic tabs are automatically updated. Points and axes may also be dragged from the biplot to the kraal.

▷ **Format...**
Opens the *Points* tab of the Format → By group dialogue box at the group of the point that was right clicked.

• **Axis pop-up menu**
Provides access to options associated with axes.

DETAIL · This pop-up menu appears when the user right clicks an axis inside the biplot. A point or axis is under the mouse if the mouse cursor changes from an arrow to a hand.

▷ **Highlight**

Highlights the axis which was right clicked.

DETAIL · This option highlights the axis which was right clicked by greying the other axes in the biplot region and diagnostic tabs. By default, the variable values of the highlighted axis are also then shown in the biplot. An axis highlight may be removed by right clicking inside the biplot and selecting Remove axis highlight from the pop-up menu. The highlight is automatically removed if the highlighted axis is moved to the kraal.

▷ **Send to kraal**

Sends the axis which was right clicked to the kraal.

DETAIL · Sending an axis to the kraal temporarily removes it from consideration. The biplot region and diagnostic tabs are automatically updated. At least three axes must be left in the biplot.

▷ **Format. . .**

Opens the Format → Axes dialogue box at the axis that was right clicked.

• **Outer pop-up menu**

Provides access to biplot options.

▷ **Show title**

The same as View → Show title.

▷ **Format title. . .**

The same as Format → Title.

▷ **Show group labels in legend**

The same as View → Show group labels in legend.

▷ **Don't show axis labels**

The same as View → Don't show axis labels.

DETAIL · Not shown when there are no axes, as opposed to the axes being hidden.

▷ **Show clinging axis labels**

The same as View → Show clinging axis labels.

DETAIL · Not shown when there are no axes, as opposed to the axes being hidden.

▷ **Show axis labels in legend**

The same as View → Show axis labels in legend.

DETAIL · Not shown when there are no axes, as opposed to the axes being hidden.

▷ **Show Additional labels in legend**

The same as View → Show Additional labels in legend.

▷ **Show next legend entries**

The same as View → Show next legend entries.

▷ **Show previous legend entries**

The same as View → Show previous legend entries.

▷ **Save as**

The same as File → Save as.

▷ **Copy**

The same as File → Copy.

▷ **Print. . .**

The same as File → Print.

▶ **Settings box**

Lets the user choose the action of the biplot axes, and effect data transformations.

DETAIL · The settings box is located towards the centre right of the GUI.

SCREENSHOT · 2.1

- **Action**

  Lets the user choose the action of the biplot axes.

  DETAIL · Disabled when there are no axes, or the axes are hidden.

  ▷ **Predict**

  A predictive biplot is shown.

  DETAIL · Predictive biplot axes are positioned and calibrated to optimally predict variable values from the positions of points in a biplot.

  ALTERNATIVE TO · Interpolate: centroid, Interpolate: vector sum

  REFERENCES · (Gower and Hand, 1996, Sections 2.3, 6.3.2)

  ▷ **Interpolate: centroid**

  A centroid interpolative biplot is shown.

  DETAIL · Centroid interpolative biplot axes are positioned and calibrated to allow for interpolation of new samples onto an existing biplot by finding a centroid.

  ALTERNATIVE TO · Predict, Interpolate: vector sum

  REFERENCES · Gower and Hand (1996, Section 5.3)

  ▷ **Interpolate: vector sum**

  A vector sum interpolative biplot is shown.

  DETAIL · Vector sum interpolative biplot axes are positioned and calibrated to allow for interpolation of new samples onto an existing biplot by finding a vector sum.

  ALTERNATIVE TO · Predict, Interpolate: centroid

  REFERENCES · (Gower and Hand, 1996, Sections 2.3, 6.3.1)

- **Transformation**

  Lets the user effect data transformations.

  DETAIL · These transformations are performed on `Data` argument of the `Biplots` function. Data are always centred to have zero column means. Log-transformations are only available if all the non-kraal variable values of the non-kraal points are strictly positive. Irrespective of the transformation, the biplot axes are always calibrated in terms of the original variable values. In order to have a biplot calibrated in terms of transformed variable values, the transformation has to be performed outside of the GUI and passed to the GUI via the `Data` argument.

  ▷ **Centre**

  Transforms the data to have zero column means.

  DETAIL · The mean of each column is subtracted from the elements of that column. The data are always centred to have zero column means.

  ALTERNATIVE TO · Centre, scale; Unitise, centre; Log, centre; Log, centre, scale; Log, unitise, centre

  ▷ **Centre, scale**

  Transforms the data to have zero column means, and then scales the data so that each column has unit variance.

  DETAIL · The mean of each column is first substracted from the elements of that column. Each column is then divided by its standard deviation.

  ALTERNATIVE TO · Centre; Unitise, centre; Log, centre; Log, centre, scale; Log, unitise, centre

▷ **Unitise, centre**

Shifts and scales each column to have minimum value 0 and maximum value 1; thereafter centres to have zero column means.

DETAIL · The minimum of each column is subtracted from the elements of that column, the answers divided by the range of the (original) column.

ALTERNATIVE TO · Centre; Centre, scale; Log, centre; Log, centre, scale; Log, unitise, centre

▷ **Log, centre**

Takes natural logarithms, then transforms the data to have zero column means.

DETAIL · Natural logarithms are taken of all elements. Thereafter the mean of each column is subtracted from the elements of that column. Only available if all the non-kraal variable values of the non-kraal points are strictly positive.

ALTERNATIVE TO · Centre; Centre, scale; Unitise, centre; Log, centre, scale; Log, unitise, centre

▷ **Log, centre, scale**

Takes natural logarithms, transforms to have zero column means, then scales so that each column has unit variance.

DETAIL · Natural logarithms are taken of all elements. Thereafter the mean of each column is first substracted from the elements of that column. Each column is then divided by its standard deviation. Only available if all the non-kraal variable values of the non-kraal points are strictly positive.

ALTERNATIVE TO · Centre; Centre, scale; Unitise, centre; Log, centre; Log, unitise, centre

▷ **Log, unitise, centre**

Takes natural logarithms, shifts and scales each column to have minimum value 0 and maximum value 1; thereafter centres to have zero column means.

DETAIL · Natural logarithms are taken of all elements. Thereafter the minimum of each column is subtracted from the elements of that column, the answers divided by the range of the (post-log) column. Only available if all the non-kraal variable values of the non-kraal points are strictly positive.

ALTERNATIVE TO · Centre; Centre, scale; Unitise, centre; Log, centre; Log, centre, scale;

▶ **Diagnostic tabs**

Consists of tabs for diagnostic graphs, dynamic variable value predictions, and quantities for export back to R.

DETAIL · The diagnostic tabs are located towards the top right of the GUI.

Note: A known issue affects the diagnostic tabs of version 0.0-2 of the BiplotGUI package.

SCREENSHOT · 2.1

• **Convergence**

If the points are determined by MDS, a graph of stress values over iterations is shown.

DETAIL · If Live updates is checked, the graph is updated live. Only available when the points are determined by MDS.

▷ **Pop-up menu**

Makes various options available with repspect to the graph of the convergence tab.

DETAIL · The convergence pop-up menu is obtained by right clicking the graph in the convergence tab.

○ **Show title**
The same as View → Show title. However, it applies to the graph in the tab rather than to the biplot region.

○ **Format...**
The same as Format → Diagnostic tabs.

○ **Save as**
The same as File → Save as. However, it it applies to the graph in the tab rather than to the biplot region.

DETAIL · Diagnostic graphs are saved with the same proportions that would be used if the currently displayed biplot region were to be saved. This ensures that the graphs can be used alongside one another. Specifically, if the biplot region has a legend, in the saved version of diagnostic graphs, an equivalent open space will be left at the bottom of the graph. To remove the space, temporarily remove the legend from the biplot and save the diagnostic graph once more.

○ **Copy**
The same as File → Copy. However, it applies to the graph in the tab rather than to the biplot region.

DETAIL · Diagnostic graphs are copied with the same proportions that would be used if the currently displayed biplot region were to be copied. This ensures that the graphs can be used alongside one another. Specifically, if the biplot region has a legend, in the copied version of diagnostic graphs, an equivalent open space will be left at the bottom of the graph. To remove the space, temporarily remove the legend from the biplot and copy the diagnostic graph once more.

○ **Print...**
The same as File → Print. However, it applies to the graph in the tab rather than to the biplot region.

DETAIL · Diagnostic graphs are printed with the same proportions that would be used if the currently displayed biplot region were to be printed. This ensures that the graphs can be used alongside one another. Specifically, if the biplot region has a legend, in the printed version of diagnostic graphs, an equivalent open space will be left at the bottom of the graph. To remove the space, temporarily remove the legend from the biplot and print the diagnostic graph once more.

○ **External**
The same as External → As is below the biplot region. However, it applies to the graph in the tab rather than to the biplot region.

DETAIL · External diagnostic graphs have the same proportions that would be used if the currently displayed biplot were to be placed in an external window. This ensures that the graphs can be used alongside one another. Specifically, if the biplot has a legend, in the diagnostic graph will have an equivalent open space. To remove the space at the bottom of the graph, temporarily remove the legend and choose the External option for the diagnostic graph once more.

• **Points**
Shows graphs of goodness of fit of the points.

DETAIL · For PCA and CVA biplots, graphs of point predictivities are shown. The first axis shows the point predictivities in the first biplot dimension. Therefore, the closer a point is to the right of the graph, the better it is represented in the first dimension

of the biplot. The distance above the diagonal line represents the contribution of the second biplot dimension. Therefore, the second axis of the point predictivities graph shows the point predictivities of the first two biplot dimensions jointly. The closer a point is to the top of the graph, the better represented it is in the two dimensions of the biplot.

When the points in a biplot are based either on PCO or MDS, a Shepard diagram is shown in this tab. In a Shepard diagram, inter-sample dissimilarities are shown on the x-axis. On the y-axis, inter-point disparities are shown on the line, while inter-point distances are shown as points. The closer the inter-point distances are to the inter-point disparities, the better the fit. By default, the five worst-fitting point pairs are identified. A Shepard diagram is only shown if there are fewer than 250 non-kraal samples.

Not available for covariance/correlation biplots.

REFERENCES · Borg and Groenen (2005, Section 3.3), Gower *et al.* (2008)

> ▷ **Pop-up menu**
> *The same as the convergence pop-up menu. However, it acts on the graphs of the points tab.*

- **Groups**
  Shows graphs of goodness of fit of the groups.

  DETAIL · For CVA biplots, graphs of group predictivities are shown. The first axis shows the group predictivities in the first biplot dimension. Therefore, the closer a point is to the right of the graph, the better represented the corresponding group is in the first dimension of the biplot. The distance above the diagonal line represents the contribution of the second biplot dimension. Therefore, the second axis of the group predictivities graph shows the group predictivities of the first two biplot dimensions jointly. The closer a point is to the top of the graph, the better represented the corresponding group is in the two dimensions of the biplot.

  Only available for CVA biplots.

  REFERENCES · Gower *et al.* (2008)

  > ▷ **Pop-up menu**
  > *The same as the convergence pop-up menu. However, it acts on the graph of the groups tab.*

- **Axes**
  Shows graphs of goodness of fit of the axes.

  DETAIL · For PCA and CVA biplots, graphs of axis predictivities are shown. The first axis shows the axis predictivities in the first biplot dimension. Therefore, the closer a point is to the right of the graph, the better represented the corresponding axis is in the first dimension of the biplot. The distance above the diagonal line represents the contribution of the second biplot dimension. Therefore, the second axis of the axis predictivities graph shows the axis predictivities of the first two biplot dimensions jointly. The closer a point is to the top of the graph, the better represented the corresponding axis is in the two dimensions of the biplot.

  Only available for PCA and CVA biplots.

  REFERENCES · Gower *et al.* (2008)

  > ▷ **Pop-up menu**
  > *The same as the convergence pop-up menu. However, it acts on the graph of the axes tab.*

- **Predictions**
  The tab in which live variable value predictions are shown.

  DETAIL · Dynamic variable value prediction is enabled by right clicking inside the biplot

and choosing either Predict cursor positions or Predict points closest to cursor positions from the pop-up menu. If Predict cursor positions is chosen, the second column of the table in the predictions tab contains the variable value predictions corresponding to the position of the mouse cursor as it moves over the biplot. If Predict points closest to cursor positions is chosen, the variable values of the point closest to the mouse cursor as it moves over the biplot, are predicted. In this case, the third column contains the actual variable values of the points. The fourth column contains the RAE, which is calculated as (prediction - actual)/(maximum actual for variable - minimum actual for variable) ×100. Biplots are not exact but optimal.

Not available when there are no axes, the axes are hidden, or the biplot is interpolative rather than predictive.

REFERENCES · Gower and Hand (1996, Sections 2.3, 6.3.2)

- **Export**
  Allows various quantities to be exported back to R to view or manipulate.

  DETAIL · Depending on the representation in the biplot region, various quantities are available for export back to R. A quantity can be displayed in the R console by selecting it and clicking Display in console. Alternatively, a quantity can be saved to the current R workspace by selecting it and clicking Save to workspace.

  REFERENCES · Gower and Hand (1996), Borg and Groenen (2005)

  - ▷ **Display in console**
    Displays the quantity selected in the export tab in the R console.

    DETAIL · Only available if a quantity has been selected in the export tab.

  - ▷ **Save to workspace**
    Saves the quantity selected in the export tab in the current R workspace under the name given in the list of quantities.

    DETAIL · Only available if a quantity has been selected in the export tab.

▶ **Kraal**
Points and axes may be dragged to the kraal, temporarily removing them from consideration.

DETAIL · The kraal is located towards the bottom right of the GUI. A point or axis can be removed to the kraal by first hovering the mouse cursor over it so that the cursor changes from an arrow to a hand, and then dragging the point or axis beyond the biplot borders. Alternatively, the point or axis can be right clicked, and Send to kraal selected from the pop-up menu. The biplot region and diagnostic tabs are then automatically updated as if the corresponding sample or variable were never part of the data set; the point or axis itself appears in the kraal. Many points and axes can simultaneously be kept in the kraal, but at least three axes must remain in the biplot.

The kraal itself is responsive to mouse clicks and motion. Different pop-up menus appear when right clicking inside the kraal, on a point, or on an axis. Points and axes may be dragged around in the kraal over a grid of possible positions, or back onto the biplot. Alternatively, the buttons below the kraal may be used to send back the points only, the axes only, or both the points and axes.

SCREENSHOT · 2.1

- **Kraal pop-up menu**
  Shows options associated with the kraal.

  DETAIL · The kraal pop-up menu is obtained by right clicking inside the kraal but not on points or axes in the kraal.

▷ **Return points**

Returns all the points in the kraal to the biplot and updates the biplot and the diagnostic tabs.

DETAIL · Not available when there are no points in the kraal.

▷ **Return axes**

Returns all the axes in the kraal to the biplot and updates the biplot and the diagnostic tabs.

DETAIL · Not available when there are no axes in the kraal.

▷ **Return all**

Returns all the points and axes in the kraal to the biplot and updates the biplot and diagnostic tabs.

DETAIL · Not available when there are no points or axes in the kraal.

▷ **Format by group...**

The same as Format → By group.

▷ **Format axes...**

The same as Format → Axes.

- **Kraal point pop-up menu**

Shows options associated with a point in the kraal.

DETAIL · The kraal point pop-up menu is obtained by right clicking a point in the kraal. When the mouse cursor is above a point or axis, the cursor changes from an arrow to a hand.

▷ **Return to biplot**

Returns the right clicked point to the biplot and updates the biplot and diagnostic tabs.

▷ **Format...**

Opens the points tab of the Format → By group dialogue box at that group of the point that was right clicked.

- **Kraal axis pop-up menu**

Shows options associated with an axis in the kraal.

DETAIL · The kraal axis pop-up menu is obtained by righ clicking an axis in the kraal. When the mouse cursor is above a point or axis, the cursor changes from an arrow to a hand.

▷ **Return to biplot**

Returns the right clicked axis to the biplot and updates the biplot and diagnostic tabs.

▷ **Format...**

Opens the Format → Axes dialogue box at that axis that was right clicked.

▶ **Other**

Contains miscellaneous features.

DETAIL · These features are located in the strip at the bottom of the GUI.

SCREENSHOT · 2.1

- **Progress bar**

Shows the progress while performing an action.

DETAIL · The progress bar appears at the bottom left of the screen while an action is being performed. While the GUI is busy, no new actions should be initiated. Changing tabs is not considered to constitute action.

- **External**
  Options to display the currently displayed biplot region in an external window.
  DETAIL · Whether external windows are re-used or not can be set by clicking File → Options.

  ▷ **As is**
  Shows the currently displayed biplot region in an external R graph.
  DETAIL · Depending on how the GUI has been resized onscreen, the proportions within an external biplot region may differ from what is shown onscreen. To ensure consistency, external biplot regions always have both their width and height set to 17.78 centimeters (7 inches). This is the default size of R graphs.
  SHORTCUT · F11

  ▷ **In 3D**
  Creates a three-dimensional version of the currently displayed biplot which can be dynamically navigated.
  DETAIL · This option is only available for non-MDS representations so that the first two dimensions of the three-dimensional biplot corresponds to the two-dimensional biplot shown in the GUI. Various options can be set by clicking File → Options. By default the left mouse button acts as a trackball for navigation, the middle mouse button allows zooming in and out, and the right mouse button sets the field of view. Not all options from the View menu or graphical parameters from the Format menu transfer to the three-dimensional biplots.
  SHORTCUT · F12
  REFERENCES · See the R help file of the `par3d` function of the `rgl` package (Adler and Murdoch, 2008) for details on the choices.

- **Hide**
  Hides the points and/or the axes of a biplot.
  ▷ **Points**
  Hides the points of the biplot.
  ▷ **Axes**
  Hides the axes of the biplot.
  DETAIL · Not available when there are no axes.

- **Live updates**
  Toggles between showing live updates of MDS convergence (both in the biplot and diagnostic tabs), and merely showing the graphs for the final, converged configuration.
  DETAIL · It can be informative to see the MDS configuration converges. However, if checked, the biplot region and Shepard diagram is redrawn after a number of iterations, taking more time. To speed up convergence, the checkbutton can be cleared. The number of iterations to the next update can be changed by clicking File → Options.

- **Stop**
  Stops the IM algorithm for MDS.
  DETAIL · Stops the iterations and proceeds with the current configuration, fitting biplot axes and all the rest. Such representations do not meet optimality criteria and may therefore be misleading. Available only while an MDS configuration is being found.

- **Return points**
  Returns all the points in the kraal to the biplot and updates the biplot and the diagnostic tabs.
  DETAIL · Not available when there are no points in the kraal.

- **Return axes**

  Returns all the axes in the kraal to the biplot and updates the biplot and the diagnostic tabs.

  DETAIL · Not available when there are no axes in the kraal.

- **Return all**

  Returns all the points and axes in the kraal to the biplot and updates the biplot and diagnostic tabs.

  DETAIL · Not available when there are no points or axes in the kraal.

Appendix A

History

A.1 Version 0.0-1

29 JULY 2008. The first release. Contains support for

- Biplots: PCA, covariance/correlation, CVA, regression, Procrustes, circular non-linear

- Interpolation and prediction

- Scaling techniques: PCO, metric MDS (identity transformation), non-metric MDS (monotone regression), semi-metric MDS (monotone spline transformation)

- Data transformations: centre; centre, scale; unitise, centre; log, centre; log, centre, scale; log, unitise, centre

- Distance metrics: Pythagoras, Square-root-of-Manhattan, Clark, Mahalanobis

- Additional descriptors: convex hulls, alpha-bags, point densities, classification regions

- Export of intermediate quantities back to R

- Diagnostic graphs: convergence; point, group and axis predictivities; Shepard diagrams

- Interactivity: variable value prediction, zooming, point and axis drag-and-drop; 3D biplots

- Customisation of graphs

- Import of data from Excel 1997–2003 files

A.2 Version 0.0-2

12 AUGUST 2008. The first CRAN release.

- The Help → Homepage option is added

- The names of two graphical parameters in the Format → Interaction dialogue box are corrected

- The `CountryData` data set is renamed the `Countries` data set

- Changes are made to the package help files, including correcting the units of the GDP variable of the countries data. The variable measures 'PPP GDP per capita in US dollars' rather than 'PPP GDP in trillions of US dollars' as previously stated. However, the countries are still those with the 15 highest PPP GDPs in US dollars during 2007 without adjusting for population size.

- The manual is updated with two new chapters included

References

ADHIKARI, K., HEYMANN, H. and HUFF, H.E. (2003). Textural characteristics of lowfat, fullfat and smoked cheeses: sensory and instrumental approaches. *Food Quality and Preference*, 14:211–218. (Cited on page 3.)

ADLER, D. and MURDOCH, D. (2008). *rgl: 3D visualization device system (OpenGL)*. R package version 0.81. Available at: http://rgl.neoscientists.org (Cited on pages 6, 13, 34, and 75.)

AGENCY, C. (2007). *The World Factbook: 2007, CIA's 2006*. Washington, DC, USA: Potomac Books. (Cited on page 7.)

ALVES, M.R., CUNHA, S.C., AMARAL, J.S., PEREIRA, J.A. and OLIVEIRA, M.B. (2005). Classification of PDO olive oils on the basis of their sterol composition by multivariate analysis. *Analytica Chimica Acta*, 549:166–178. (Cited on page 4.)

BLASIUS, J., EILERS, P. and GOWER, J.C. (2008). Better biplots. *Computational Statistics and Data Analysis*. In press. (Cited on pages 18 and 62.)

BORG, I. and GROENEN, P.J.F. (2005). *Modern Multidimensional Scaling: Theory and Applications*. Springer Series in Statistics, 2nd edition. New York, NY, USA: Springer. (Cited on pages 3, 15, 33, 49, 50, 51, 72, and 73.)

BOTES, A., MCGEOCH, M.A. and VAN RENSBURG, B.J. (2006). Elephant- and human-induced changes to dung beetle (coleoptera: Scarabaeidae) assemblages in the maputuland centre of endemism. *Biological Conservation*, 130:573–583. (Cited on page 3.)

BRADU, D. and GABRIEL, K.R. (1978). The biplot as a diagnostic tool for models of two-way tables. *Technometrics*, 20:47–68. (Cited on page 2.)

BRAUN, W.J. and MURDOCH, D.J. (2007). *A First Course in Statistical Programming with R*. Cambridge, UK: Cambridge University Press. (Cited on page 5.)

BURDEN, M., GARDNER, S., LE ROUX, N.J. and SWART, J.P.J. (2001). Ou-kaapse meubels en stinkhout-identifikasie: Moontlikhede met kanoniese veranderlike-analise en bistippings. *South African Journal of Cultural History*, 15:50–73. (Cited on page 18.)

CHAMBERS, J.M. (2007). *Software for Data Analysis: Programming with R*. New York, NY, USA: Springer. (Cited on page 5.)

CLARK, P.J. (1952). An extension of the coefficient of divergence for use with multiple characters. *Copeia*, 2:61–64. (Cited on page 49.)

COOK, R.D. and WEISBERG, S. (1982). *Residuals and influence in regression*. Monographs on Statistics and Applied Probability. London, UK: Chapman & Hall. (Cited on page 21.)

COX, T.F. and COX, M.A.A. (2001). *Multimensional Scaling*. Monographs on Statistics and Applied Probability, 2nd edition. Boca Raton, FL, USA: Chapman & Hall/CRC. (Cited on pages 3, 49, and 52.)

DAWKINS, B. (1989). Multivariane analysis of national track records. *The American Statistician*, 43(2):110–115. (Cited on page 4.)

DE LEEUW, J. (1977). Applications of convex analysis to multidimensional scaling. In J.R. BARRA, F. BRODEAU, G. ROMIER and B. VAN CUTSEM (eds.), *Recent developments in statistics*, pages 133–145. Amsterdam, The Netherlands: North-Holland Publishing Company. (Cited on page 15.)

DE LEEUW, J. and HEISER, W.J. (1980). Multidimensional scaling with restrictions on the configuration. In P.R. KRISHNAIAH (ed.), *Multivariate Analysis*, volume V, pages 501–522. Amsterdam, The Netherlands: North-Holland Publishing Company. (Cited on page 15.)

DRAY, S. and DUFOUR, A.-B. (2007). The ade4 package: Implementing the duality diagram for ecologists. *Journal of Statistical Software*, 22(4):1–20. (Cited on page 5.)

DRAY, S. and DUFOUR, A.-B. (2008). *ade4: Analysis of Ecological Data: Exploratory and Euclidean methods in Environmental sciences*. R package version 1.4-9. Available at: http://pbil.univ-lyon1.fr/ade4 (Cited on page 5.)

GABRIEL, K.R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika*, 58(3):453–467. (Cited on pages 1, 2, 5, 48, and 66.)

GABRIEL, K.R. (1972). Analysis of meteorological data by means of canonical decomposition and biplots. *Journal of Applied Meteorology*, 11:1071–1077. (Cited on page 2.)

GABRIEL, K.R. and MATHER, G.K. (1986). Exploratory data analysis of 1951–82 summer rainfall around nelspruit, transvaal, and possible effects of 1972–81 cloud seeding. *Journal of Climate and Applied Meteorology*, 25(8). (Cited on page 4.)

GARDNER, S. (2001). Extensions of biplot methodology to discriminant analysis with applications of non-parametric princial components. Unpublished PhD thesis, Stellenbosch University, Stellenbosch, South Africa. (Cited on pages 3, 5, 20, 48, and 59.)

GARDNER, S. and LE ROUX, N.J. (2005). Extentions of biplot methodology to discriminant analysis. *Journal of Classification*, 22:59–86. (Cited on pages 20 and 63.)

GARDNER, S. and LE ROUX, N.J. (2006). Sub-species of homopus areolatus? Biplots and small class inference with analysis of distance. In V. BATAGELJ, H.-H. BOCK, A. FERLIGOJ and A. ZIBERNA (eds.), *Data Science and Classification*, pages 221–228. Proceedings of the 10th Jubilee Conference of the International Federation of Classification Societies, Berlin, Germany: Springer. (Cited on page 4.)

GARDNER, S., LE ROUX, N.J., RYPSTRA, T. and SWART, J.P.J. (2005). Extending a scatterplot for displaying group stucture in multivariate data: A case study. *ORiON*, 21(2):111–124. (Cited on pages 3 and 13.)

GOWER, J.C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53(3/4):325–338. (Cited on pages 3 and 49.)

GOWER, J.C. (1990). Three dimensional biplots. *Biometrika*, 77:773–785. (Cited on page 2.)

GOWER, J.C. (1995). A general theory of biplots. In W.J. KRZANOWSKI (ed.), *Recent Advances in Descriptive Multivariate Analysis*, pages 283–303. Oxford, UK: Oxford University Press. (Cited on pages 2 and 3.)

GOWER, J.C. and DIJKSTERHUIS, G.B. (2004). *Procrustes Problems*. Oxford Statistical Science Series. Oxford, UK: Oxford University Press. (Cited on pages 33 and 52.)

GOWER, J.C. and HAND, D.J. (1996). *Biplots*. Monographs on Statistics and Applied Probability. London, UK: Chapman & Hall. (Cited on pages 2, 3, 5, 21, 22, 23, 33, 35, 48, 49, 52, 55, 63, 66, 67, 69, and 73.)

GOWER, J.C. and HARDING, S.A. (1988). Nonlinear biplots. *Biometrika*, 75(3):445–455. (Cited on pages 3, 49, and 52.)

GOWER, J.C. and KRZANOWSKI, W.J. (1999). Analysis of distance for structured multivariate data and extensions to multivariate analysis of variance. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 48(4):505–519. (Cited on page 3.)

GOWER, J.C., LE ROUX, N.J. and GARDNER-LUBBE, S. (2008). Measures of fit in principal component and canonical variate analyses. *Journal of Applied Statistics*, 35(9):947–965. (Cited on pages 10, 13, 21, and 72.)

GOWER, J.C. and LEGENDRE, P. (1986). Metric and euclidean properties of dissimilarity coefficients. *Journal of Classification*, 16:5–48. (Cited on page 48.)

GOWER, J.C. and NGOUENET, R.F. (2005). Nonlinearity effects in multidimensional scaling. *Journal of Multivariate Analysis*, 94:344–365. (Cited on pages 3, 49, and 52.)

GRAFFELMAN, J. (2005). Enriched biplots for canonical analysis. *Journal for Applied Statistics*, 32(2):173–188. (Cited on page 2.)

GRAFFELMAN, J. (2007). *calibrate: Calibration of Biplot Axes*. R package version 1.4. (Cited on page 5.)

GRAFFELMAN, J. and VAN EEUWIJK, F.A. (2005). Calibration of multivariate scatter plots for exploratory analysis of relations within and between sets of variables in genomic research. *Biometrical Journal*, 47(6):863–879. (Cited on page 5.)

GREENACRE, M.J. (1984). *Theory and Applications of Correspondence Analysis*. London, UK: Academic Press. (Cited on pages 2 and 48.)

GREENACRE, M.J. (2007). *Correspondence Analysis in Practice*. Interdisciplinary Statistics, 2nd edition. London, UK: Chapman & Hall/CRC. (Cited on page 4.)

GROSJEAN, P. (2008). *tcltk2: Additional Tcl/Tk widgets and commands for R*. R package version 1.0-7. Available at: http://www.sciviews.org/SciViews-R (Cited on page 6.)

HABER, M. and GABRIEL, K.R. (1976). Weighted least squares approximation of matrices and its application to canonical correlations and biplot display. Technical Report, University of Rochester, Department of Statistics. (Cited on page 2.)

HIGHLAND STATISTICS LTD (2007). brodgar. Version 2.5.5. Available at: http://www.brodgar.com/brodgar.htm (Cited on page 4.)

HOFMANN, H. (2000). Manet. Version 1862. Available at: http://stats.math.uni-augsburg.de/manet (Cited on page 4.)

HOL, P.J., MOLSTER, A. and GJERDET, N.R. (2008). Should the galvanic combination of titanium and stainless steel surgical implants be avoided? *Injury*, 39:161–169. (Cited on page 3.)

HOTELLING, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24:417–441. (Cited on page 3.)

HOTELLING, H. (1935). The most predictable criterion. *Journal of Educational Psychology*, 26:139–142. (Cited on page 3.)

HOTELLING, H. (1936). Relations between two sets of variables. *Biometrika*, 28:321–377. (Cited on page 3.)

IHAKA, R. and GENTLEMAN, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3):299–314. (Cited on page 4.)

JEMWA, G.T. and ALDRICH, C. (2006). Kernel-based fault diagnosis on mineral processing plants. *Minerals Engineering*, 19:1149–1162. (Cited on page 4.)

KOVACH COMPUTING SERVICES (2008). mvsp. Version 3.1. Available at: http://www.kovcomp.co.uk/mvsp (Cited on page 4.)

Kruskal, J.B. (1964). Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29:115–129. (Cited on page 15.)

Kruskal, J.B. and Wish, M.W. (1978). *Multidimensional Scaling*. Quantitative Applications in the Social Sciences. Thousand Oaks, CA, USA: Sage Publications. (Cited on page 3.)

Krzanowski, W.J. (2000). *Principles of Multivariate Analysis: A User's Perspective*. Oxford Statistical Science Series, revised edition. New York, NY, USA: Oxford University Press. (Cited on page 18.)

Krzanowski, W.J. (2004). Biplots for multifactorial analysis of distance. *Biometrics*, 60:517–524. (Cited on pages 3 and 13.)

Le Roux, N.J. and Gardner, S. (2005). Analysing your multivariate data as a pictorial: A case for applying biplot methodology? *International Statistical Review*, 73(3):365–387. (Cited on pages 4 and 18.)

Lipkovich, I. and Smith, E.P. (2002a). Biplot. Available at: http://filebox.vt.edu/artsci/stats/vining/keying/biplot_final.zip (Cited on page 4.)

Lipkovich, I. and Smith, E.P. (2002b). Biplot and singular value decomposition macros for Excel. *Journal of Statistical Software*, 7(5):1–15. (Cited on page 4.)

Loomis, E.S. (1968). *The Pythagorean Proposition*. 2nd edition. Washington, DC, USA: National Council of Teachers of Mathematics. (Cited on page 49.)

Losa, F.B., van den Honert, R. and Joubert, A. (2001). The multivariate analysis biplot as a tool for conflict analysis in MCDA. *Journal of Multi-Criteria Decision Analysis*, 10:273–284. (Cited on page 3.)

Mahalanobis, P.C. (1936). On the generalised distance in statistics. *Proceedings of the National Institute of Science of India*, 12:49–55. (Cited on pages 2 and 49.)

Meyer, D., Zeileis, A. and Hornik, K. (2008). *vcd: Visualizing Categorical Data*. R package version 1.0-9. (Cited on page 6.)

Minitab Inc (2007). Minitab. Version 15. Available at: http://www.minitab.com (Cited on page 4.)

MjM Software Design (2007). Pc-ord. Version 5.12. Available at: http://home.centurytel.net/~mjm/pcordwin.htm (Cited on page 4.)

Munday, P.L., Jones, G.P. and Caley, M.J. (1997). Habitat specialisation and the distribution and abundance of coral-dwelling gobies. *Marine Ecology Progress Series*, 152:227–239. (Cited on page 3.)

Murrel, P. (2005). *R Graphics*. Boca Raton, FL, USA: Chapman & Hall/CRC. (Cited on page 5.)

Naidoo, S., Harris, A., Swanevelder, S. and Lombard, C. (2006). Fetal alcohol syndrome: A cephalometric analysis of patients and controls. *European Journal of Orthodontics*, 28:254–261. (Cited on page 4.)

Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Simpson, G.L., Stevens, M.H.H. and Wagner, H. (2008). *vegan: Community Ecology Package*. R package version 1.13-1. Available at: http://vegan.r-forge.r-project.org/ (Cited on page 5.)

Pearson, K. (1901). On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2:559–572. (Cited on page 3.)

Plant Research International (2002). Canoco. Version 4.5. Available at: http://www.pri.wur.nl/uk/products/canoco (Cited on page 4.)

R Development Core Team (2008). *R: Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. Available at: http://www.R-project.org (Cited on pages 4 and 6.)

RAMSEY, J.O. (1982). Some statistical approaches to multidimensional scaling data. *Journal of the Royal Statistical Society. Series A (General)*, 145(3):285–312. (Cited on page 15.)

RAMSEY, J.O. (1988). Monotone regression splines in action. *Statistical Science*, 3(4):425–441. (Cited on page 15.)

ROUSSEEUW, P.J., RUTS, I. and TUKEY, J.W. (1999). The bagplot: A bivariate boxplot. *The American Statistician*, 53(4):382–387. (Cited on pages 21 and 59.)

SAS (2008). Available at: http://www.sas.com (Cited on page 4.)

SCHOENBERG, I.J. (1935). Remarks to Murice Fréchet's article "Sur la définition axiomatique d'une classe d'espaces vectoriels distanciés applicables vectoriellement sur l'espace de Hilbert.". *Annals of Mathematics*, 36:724–732. (Cited on page 3.)

SHEPARD, R.N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. *Psychometrika*, 27(2):125–140. (Cited on page 15.)

SINGH, R.P., HUERTA-ESPINO, J., SHARMA, R., JOSHI, A.K. and TRETHOWAN, R. (2007). High yielding spring bread wheat germplasm for global irrigated and rainfed production systems. *Euphytica*, 157:351–363. (Cited on page 3.)

SMILAUER, P. (2003). Canodraw. Version 4.1. Available at: http://www.canodraw.com (Cited on page 4.)

SPECTOR, P. (2008). *Data Manipulation with R*. New York, NY, USA: Springer. (Cited on page 5.)

SPSS INC (2007). Spss. Version 16. Available at: http://www.spss.com (Cited on page 4.)

STANLEY, W. and MILLER, M. (1979). Measuring technological change in jet fighter aircraft. Technical Report R-2249-AF, RAND Corporation, Santa Monica, CA, USA. (Cited on page 21.)

STATACORP LP (2007). Stata. Version 10. Available at: http://www.stata.com (Cited on page 4.)

STATSOFT INC (2007). Statistica. Version 8. Available at: http://www.statsoft.com (Cited on page 4.)

SUTER, H.-P., TREETRON and SWITZERLAND (2006). *xlsReadWrite: Natively read and write Excel files*. R package version 1.3.2. Available at: http://treetron.googlepages.com (Cited on page 6.)

TER BRAAK, C.J.F. (1990). Interpreting canonical correlation analysis through biplots of structure correlations and weights. *Psychometrika*, 55(3):519–531. (Cited on page 2.)

TER BRAAK, C.J.F. and LOOMAN, C.W.N. (1994). Biplots in reduced-rank regression. *Biometrical Journal*, 36(8):983–1003. (Cited on page 2.)

THIOULOUSE, J. and DRAY, S. (2007a). *ade4TkGUI: ade4 Tcl/Tk Graphical User Interface*. R package version 0.2-2. Available at: http://pbil.univ-lyon1.fr/ade4 (Cited on page 5.)

THIOULOUSE, J. and DRAY, S. (2007b). Interactive multivariate data analysis in R with the ade4 and ade4tkgui packages. *Journal of Statistical Software*, 22(5):1–14. (Cited on page 5.)

TIERNEY, L. (1990). *Lisp-Stat: An Object-Oriented Environment for Statistical and Dynamic Graphics*. Wiley Series in Probability and Mathematical Statistics. New York, NY, USA: Wiley. (Cited on page 4.)

TIERNEY, L. (2008). *tkrplot: TK Rplot*. R package version 0.0-18. (Cited on page 6.)

UDINA, F. (2005a). Interactive biplot construction. *Journal of Statistical Software*, 13(5):1–16. (Cited on page 4.)

UDINA, F. (2005b). Xls-biplot. Version 1.1a. Available at: http://tukey.upf.es/xls-biplot/index.html (Cited on page 4.)

UNDERHILL, L.G. (1990). The coefficient of variation biplot. *Journal of Classification*, 7:241–256. (Cited on pages 2 and 4.)

VANOLI, M., ZERBINI, P.E., GRASSI, M., RIZZOLO, A., FORNI, E., CUBEDDU, R., PIFFERI, A., SPINELLI, L. and TORRICELLI, A. (2006). Pectic composition, optical properties measured by time-resolved reflectance spectroscopy and quality in 'jonagored' apples. *Journal of Fruit and Ornamental Plant Research*, 14:273–282. (Cited on page 3.)

VENABLES, W.N. and RIPLEY, B.D. (2002). *Modern Applied Statistics with S*. Statistics and Computing, 4th edition. New York, NY, USA: Springer. Available at: http://www.stats.ox.ac.uk/pub/MASS4 (Cited on page 6.)

VSN INTERNATIONAL LTD (2007). Genstat. Version 10.1. Available at: http://www.genstat.com (Cited on page 4.)

WAND, M. (2008). *KernSmooth: Functions for kernel smoothing for Wand and Jones (1995)*. Ported to R by Ripley BD. R package version 2.22-22. (Cited on pages 6, 18, and 62.)

WRC RESEARCH SYSTEMS INC (2007). Brandmap. Version 7. Available at: http://www.wrcresearch.com/brandmap50.htm (Cited on page 4.)

WURZ, S., LE ROUX, N.J., GARDNER, S. and DEACON, H.J. (2003). Discriminating between the end products of the earlier middle stone age sub-stages at klasies river using biplot methodology. *Journal of Archaeological Science*, 30:1107–1126. (Cited on pages 20 and 59.)

YAN, W. and KANG, M.S. (2003). *GGE Biplot Analysis: A Graphical Tool for Breeders, Geneticists, and Agronomists*. Boca Raton, FL, USA: CRC Press. (Cited on pages 3 and 4.)

YAN, W. and KANG, M.S. (2006). Ggebiplot. Version 5. Available at: http://www.ggebiplot.com (Cited on page 4.)

YOUNG, F.W. (2001). Vista. Version 6.4. Available at: http://visualstats.org (Cited on page 4.)