RR 88-04

GRAPHICAL TOOLS IN MULTIVARIATE ANALYSIS

Peter Verboon

Department of Data Theory

Department of Human and Computer Studies

University of Leiden

Introduction

In this paper a short outline is given of some graphical applications within the context of multivariate data analysis. A first use of graphical methods is in the presentation of conclusions. The graphics used for this purpose should be simple and straightforward, because they need to communicate results to a fairly large audience. The classical histogram is an example of this type of graphic. Another use of graphical methods is in data analysis. The data analyst uses the graphical tools to answer questions about the data. Sometimes questions might be invoked by a particular graphic, which may lead the analyst to start new analyses. For an historical outline of graphical methods in statistics see Cox(1978) and Wainer and Thissen (1981). In this paper some graphical methods for multivariate data analysis are discussed. The discussion is restricted to static graphics, so graphical techniques for dynamic data analysis are not concerned. First a general introduction into the field of multivariate analysis will be given.

Multivariate data are normally collected in the form of a data matrix, with n rows and m columns. The columns usually represent the variates or variables. The rows then are the objects of study, which contribute values to the variables. This can be anything interesting enough to study: persons, animals, cars, countries etc. The rows of the data matrix will be referred to as objects, the most general term.

It should be realized though that the variables can be important objects of study too. Graphical methods for data analysis should therefore account for both aspects of the multivariate data: the variables and the objects.

There are many definitions of multivariate analysis (see Gifi, 1981) but for now it is sufficient to state that the basic goal of multivariate analysis is to discover structure in the data matrix. There are a lot of techniques however to reach this goal. Which of these techniques is appropriate for analyzing multivariate data depends on several factors.

In the first place we have factors directly related to the data. Sometimes variables may be grouped into sets according to some criterion. The different sets may then be symmetrical or asymmetrical, which means that the variables from one set may or may not depend on the variables from the other set. Another important factor is the measurement level of the variables. There are continuous and discrete variables. Discrete variables can be numerical, ordinal or nominal, with each level eliciting its own class of techniques. Furthermore there might be explaining variables and explained variables.

There are also factors that are independent of the data. For example the tradition in which the data analyst has been educated will guide him to some technique, while keeping him off others. Finally the specific goals of the data analyst will be an important factor to decide upon which technique to use. He may for instance want some variables to be explained by some specific others. Although this may sound somewhat dubious, it is clear that researchers analyzing their data to confirm or reject some model will probably be using other techniques than those without any models on their mind.

This brings us to the more general distinction within multivariate analysis, namely the distinction between confirmatory and exploratory data analysis .

Having some more or less specific model about 'reality' in mind and using the data to confirm or reject this model, we are doing confirmatory data analysis. In other words we try to fit a model on the data. We let the data speak, but restrict its answer to yes or no only. In fact this means that assumptions about the distributions of the variables are necessary.

On the other hand in exploratory data analysis (see Hartwig & Dearing, 1979) we try to learn as much as possible from the data and finally a model is extracted from the data. We let the data speak frankly. No assumptions whatsoever are needed.

From now I will concentrate on exploratory data analysis (EDA), because in this type of analysis the application of graphics is already wide-spread and by far the most important tool.

The basic thought of EDA is division of the data into two parts, an underlying structured part (smooth) and a random or error part (rough). Tukey (1977) writes this down as :

data = smooth + rough

The aim of EDA is to extract as much smooth as possible from the rough, until 'the rough is rough enough'. This implicates another important feature, the iterative process of EDA. An analysis is done and examining the outcomes new analyses follow. The major tools to establish this process are graphics. To learn the usefulness of graphics in exploratory multivariate analysis, some graphical tools for univariate and bivariate data will be examined first.

Univariate data

With only one variable in the data set, the interesting object of study is the distribution of this variable. There are three basic aspect of a distribution, the first one is the location, which is usually represented by the mean, though other statistics might sometimes do better. The median for instance is a more robust statistic when the distribution contains some outliers. A second aspect is the spread of the distribution. For this aspect the standard deviation is often computed, although this also lacks the quality of robustness when outliers are involved. Finally the shape of the distribution is an important feature of a variable. A shape can be symmetrical or skewed, single- or multipeaked and so on.

To display the aspects of a distribution, we have several graphical means at our disposal. The most important ones will now be briefly discussed.

Stem-and-leaf diagrams

A stem-and-leaf diagram, originally introduced by Tukey (1977), is particularly useful when dealing with small data sets. To create one, the observed values must be separated into a stem, the first one or first few digits, and a leaf, the remaining part of the data value. The stem values are listed in descending or ascending order and separated from their matching leafs with a vertical line. In Figure 1 an example is shown.

0	1223
1	0145
2	34457889
3	01122235677899
4	00112234456
5	13578
6	248
7	
8	3
9	4

Figure 1. Example of stem-and-leaf diagram.

The stem-and-leaf diagram can actually be looked upon as a labeled barchart, from which the basic aspects of a distribution can be read. We easily obtain a good impression of the shape: peaks, gaps and outliers are detected just by looking. The improvements over the classical histogram are the easier way of making a stem-and-leaf diagram by hand and the fact that it diminishes the arbitrariness of choosing the bars.

Box-plots

A more visual representation is the box-plot, also developed by Tukey (1977). See Figure 2.

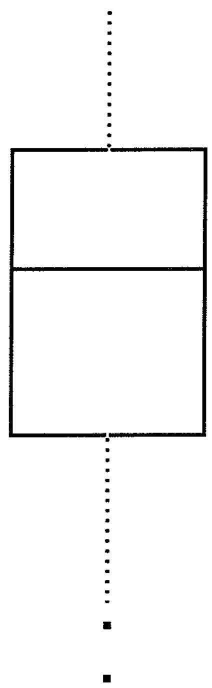

Figure 2. Example of Box-plot.

The upper and lower quartiles of the data are portrayed by the top and bottom of the rectangle. The median is the horizontal line within the rectangle. The difference between the two quartiles is called the interquartile range (IQR). The dashed lines (whiskers) extend in both sides to the largest and smallest values of the data, provided that these are within the range of 1.5 times the IQR added to the corresponding quartile value. Values outside this range are portrayed as individual points. The box-plot also gives a good rough indication of the shape of a distribution and can be used with large data sets. Outliers are easily detected.

Quantile-Quantile plots

An interesting way to compare the distribution of a variable with some other variable or with a theoretical distribution is to make a quantile-quantile (Q-Q) plot (Wilk & Gnanadesikan, 1968; Gnanadesikan, 1977). The only thing to do is to order the data and to compute quantiles. These quantiles can be plotted against the corresponding quantiles of another variable or against for example z-scores. If, after plotting, the straight line y = x appears, the shape of the two distributions is the same. In Figure 3 a Q-Q plot is shown for a variable with a distribution that is slightly skewed to the right.

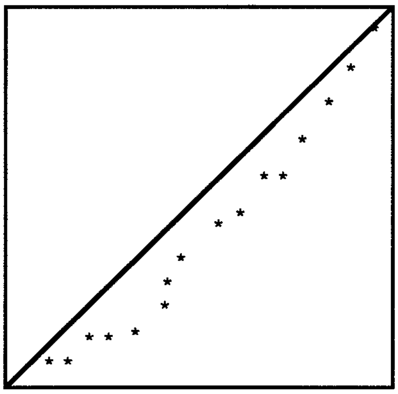

Figure 3. Example of Q-Q plot

In this plot the horizontal axis represents the quantiles of the normal distribution, while the vertical axis represents the quantiles of the variable of interest. We may conclude from this plot that the variable is not normally distributed, because the points do not fall on the line y = x. Furthermore it can be seen that all points are positioned below the line y=x, and that the density is higher at the lower side of the range of the variable. The variable is said to be positively skewed.

Bivariate data

With two variables involved the main interest is in the relationship between these variables. A linear relationship is usually the first to be examined. This is done by computing the Pearson product-moment correlation, r. We should realize though that a specific value of r may arise from quite different types of relationships (see also Chambers et al 1983; Tukey, 1986). Therefore a scatter plot should be made, which plots the value of the first variable against the corresponding value of the second for each object. Now the nature of r can be better understood.

In Figure 4 an example is shown of different relationships, that all result, however, in approximately the same value for r. The example clearly shows the usefulness of making scatter plots for examining relationships between two variables.

When there are many points the plot might become confusing, while for instance in some regions of the picture the density of the points becomes too high. Some methods are proposed to cope with this problem, for example drawing circles of varying size, corresponding with the density. The use of symbols for different densities is another way to tackle this problem (Chambers et al, 1983).

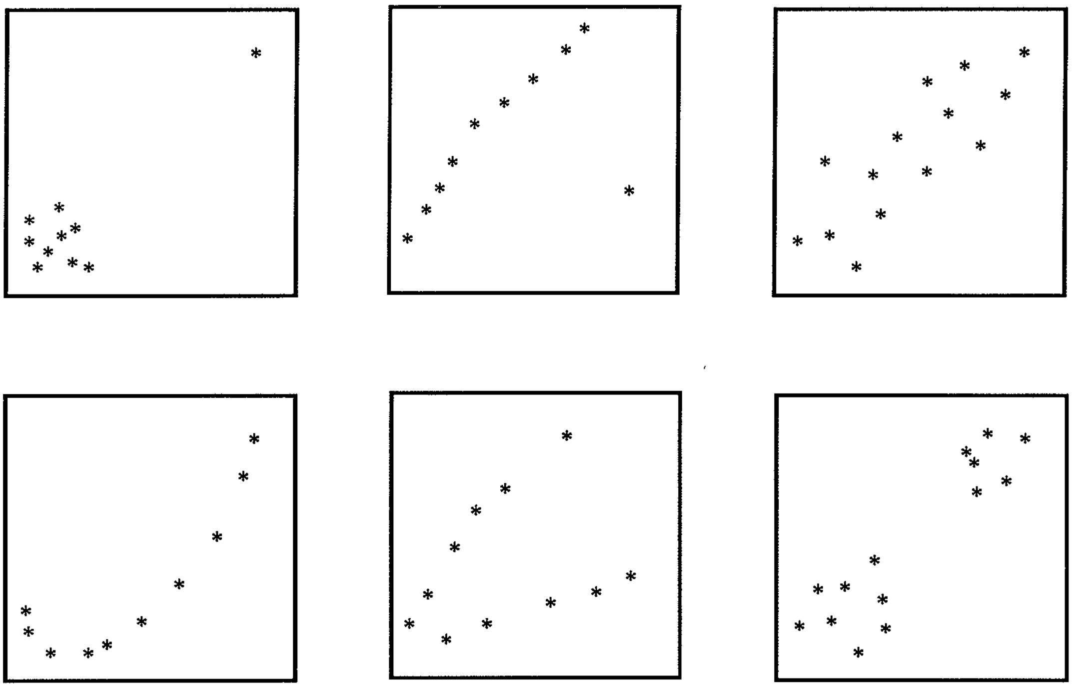

Figure 4. Example of scatter plots with equal correlation coefficient

Regression Analysis

In EDA terminology a relationship can be sharpened by smoothing it. The well known method of least squares regression is an example of this. By this method we obtain not only a straight line, representing the linear relationship, but also a plot of deviations (the rough) from the straight line, which is usually called the residual part. This plot might also reveal some interesting patterns within the data. For instance see Figure 5.

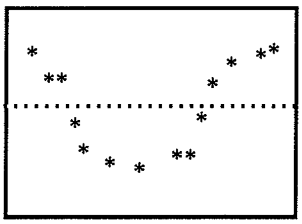

Figure 5. Residual plot

A pattern like in Figure 5 is typical when the relationship is quadratic.

Although clear and straightforward pictures like Figure 5 probably won't emerge all too often, plotting of residuals is an important way to gain some insight into the structure of the data.

Correspondence Analysis

When the two variables are both discrete with a nominal measurement level, the data matrix is usually presented in the form of a contingency table. This table can be analyzed by correspondence analysis (Greenacre, 1984). Correspondence analysis provides a graphical display containing points that represent the rows and columns of the contingency table. This technique gives insight into the relation between the separate categories of the variables. Nothing is learned with respect to the objects however.

Multivariate data

When there are more than two variables, the objective is to find structure within this set of variables. We cannot do with one simple plot, however. More complicated ways are needed to show us relations within the data. Basically there are two approaches. The first is to show the high-dimensional space containing all the variables. A lot of ingenious methods have been developed in the last decades for this goal. See also Wainer and Thissen (1981), Chambers and Kleiner (1982) and Wainer (1983).

Quite a different approach is dimension reduction. This is best explained geometrically. The n objects can be imagined as n vectors in the m-dimensional space of the variables (see Van de Geer, 1986). With only one, two or three variables this is quite easy to visualize, but with more variables the same principle holds. On the other hand the m variables can be imagined as m vectors in the n-dimensional space of the objects.

To reduce the dimensionality the objects in the m-dimensional space of the variables are projected onto a low-dimensional space. This space is mostly two-dimensional, because of the possibility to depict the objects in only one display, which increases the interpretability. This is however not the only criterion for choosing the number of dimensions in the low-dimensional space. More general the low-dimensional space is computed in such a way that the loss of information is as small as possible. This broad rule can be elaborated into several mathematical loss functions (see Meulman, 1986).

We will now start by discussing some of the possibilities for representing the objects in the high-dimensional space of all the variables. The graphical techniques used for this goal can roughly be distinguished in techniques using symbolic representations and techniques that are extensions of the ordinary scatter plot.

Symbolic representations

Polygons

The idea of using polygons (and also glyphs) was suggested by Anderson (1960). The variables are depicted as rays along equally spaced radii from a common center. The length of a ray represents the value of the variable. Connecting the end points of the rays results in an irregularly shaped polygon. For each object a separate polygon is made. See for an example with m=8 and n=4 Figure 6. Objects with approximately the same pattern can rather easily be detected. Therefore this method is particularly useful when the interest is in the arrangement of the objects.

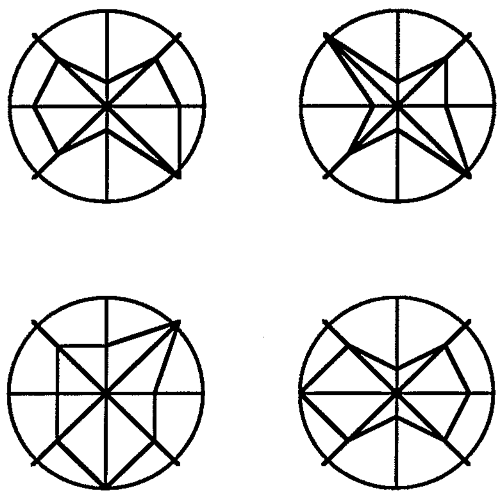

Figure 6. Example of polygons.

The shortcomings of this method are obvious: with more than say eight variables the display becomes rather illegible. Furthermore because the order of the variables is arbitrary, the shape of the polygon is arbitrary too. And finally with a lot of objects the method also becomes difficult to read.

When the number of objects is small, there is of course no problem in displaying the objects as rays of a polygon. This may enable us to find clusters of correlated variables. Variables that are not correlated to each other should be represented by very different polygons.

Profiles

A method related to the polygons is the use of profiles. Each object is represented by *m* vertical bars. The length of each bar is proportional to the value of the corresponding variable. The top of the bars can be connected by a line. The same shortcomings hold as with polygons.

Andrews' curves

Andrews (1972) proposed a method that involves the calculation of a periodic function of *m*-Fourier components for each object. This results in a curve which is a linear combination of the variables. All curves are displayed in one plot. Curves with approximately the same shape refer to objects with similar patterns. The problem of too many variables has been solved with this method, but an additional problem is that the number of objects should even be more restricted than with polygons and profiles.

Inside-out plots

When all variables in a data matrix have been standardized, the data can be plotted "inside-out" (Ramsay, 1980). Although the result of this method can hardly be called a plot, it is still an interesting tool. The standardized values are placed as the left column of a table (outside) and the object labels for each variable in the resulting *m* columns (inside). For each variable we can see which objects are deviant, and furthermore we get an overall picture of the data.

Faces

Another method is the use of faces, originally developed by Chernoff (1973). The size, shape or orientation of each feature of a face is used to represent a variable. Thus the size of the eyes corresponds with the value of a particular variable and the length of the mouth with another. Computer programs have been developed to draw faces representing up to 18 variables. Jacob et al (1976) and Wainer (1979) have shown that the application of faces can be useful in clustering the objects.

Polyhedron Graphs

Turner et al (1986) introduced the Polyhedron graph, which is a generalization of the polygon. This graph is especially useful when the data set contains one or two variables that can be considered as dependent variables. For example in a multiple regression problem with one dependent variable and m-1 independent variables, the independent variables constitute the basic polygon and the dependent variable the vertex outside this base. Contrary to the ordinary polygon which was discussed above, the base of the polyhedron is positioned only on the front half of a circle.

Figure 7 shows an example of the polyhedron graph with five independent variables and one dependent variable. The dependent variable is represented by the height of the polyhedron. With two independent variables a line can be drawn downwards from the center of the imaginary circle.

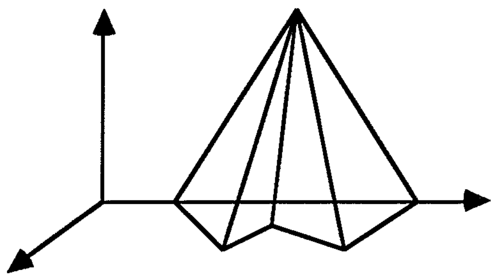

Figure 7. Example of polyhedron

This techniques also lacks the power to handle many variables or many objects. It does show, however, some aspects of the relation between dependent and independent variables.

Trees

The methods discussed so far might be useful for finding clusters of objects. If, however, we are also interested in the clustering of the variables these methods are not appropriate. Kleiner and Hartigan (1981) have introduced a rather simple technique for this problem: the tree.

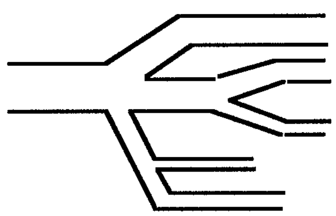

Figure 8. Example of tree plot

An hierarchical clustering of the variables is used to structure the branches of the tree. Variables with a high correlation are depicted as branches close to each other. The resulting tree is used as basic symbol. The length of each branch varies per object, representing the value of a variable. An example is shown in Figure 8.

Extensions of the scatter plot

When the number of variables is small there are some extensions of the ordinary scatter plot that can be used for exploring the data. For the following examples the simple data from table 1 will be used. This data set consists of four variables and nine objects.

Table 1.

A	B	C	D
1	2	1	1
2	4	1	1
3	6	1	1
4	8	2	1
5	6	2	2
6	4	2	2
7	5	3	2
8	7	3	2
9	9	3	2

Multiple scatter plots

A multiple scatter plot, also called (generalized) draftsman's display, is a straightforward extension of the ordinary scatter plot. Each variable is systematically plotted against every other variable. This results in a display containing $m(m\text{-}1)/2$ separate scatter plots. The adjacent plots always have one axis in common. This helps in comparing the plots. See Figure 9.

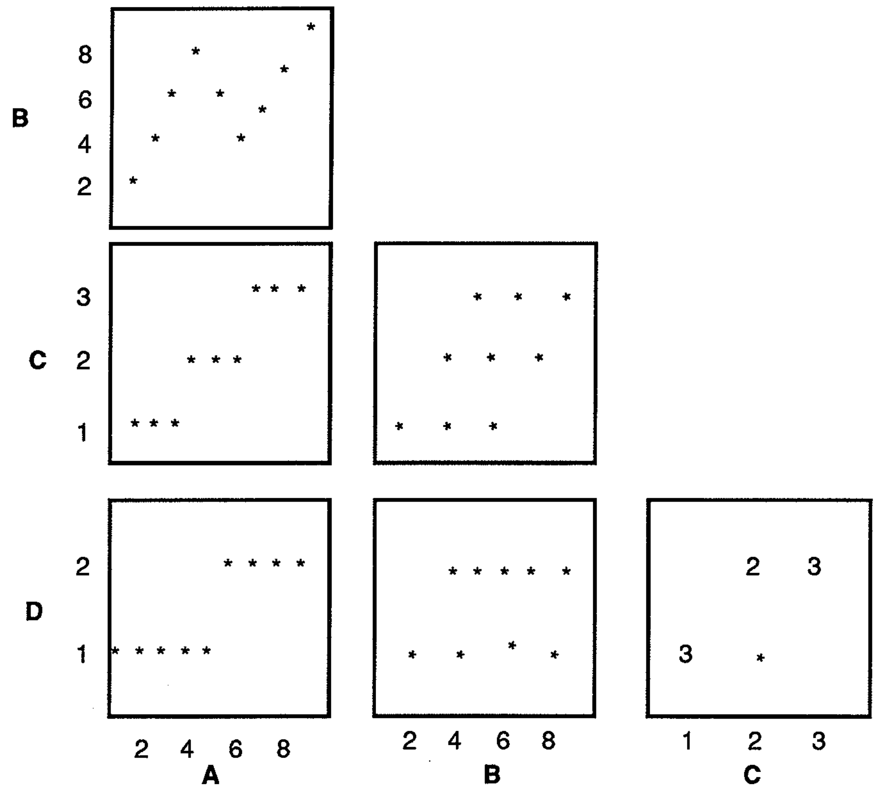

Figure 9. Example of Multiple Scatter plot

Casement displays

A casement display (Tukey and Tukey, 1983) is best explained by visualizing the values of three variables in a cube. If we take slices of this cube and place them side by side, we obtain a casement display. So for some category or some interval of a variable, the objects that are in this category or interval are plotted against each other according to the values of two other variables. See Figure 10. This figure shows clearly that the relation between **A** and **B** depends on the value of **C**. This is called second order interaction. If there exist a possibility to use colors for the graphs and if the number of objects is not too large, plots like the one in Figure 10 can be joined into one plot using different colors for each level of the third variable. In such a plot crossing lines indicate the presence of interaction.

Notice that in the casement display we can only show the relation between three variables at the same time.

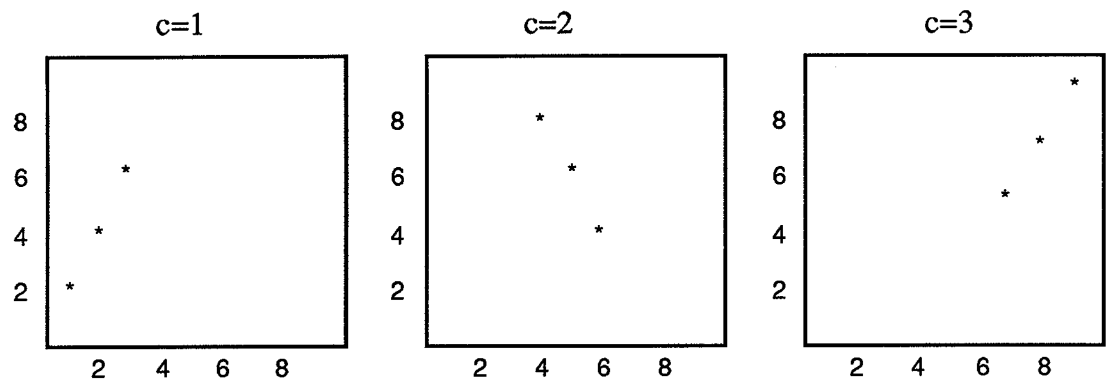

Figure 10. Example of Casement display

Multiwindow displays

If we extend a casement display to four variables, we have a multiwindow display (Chambers et al ,1983). See Figure 11.

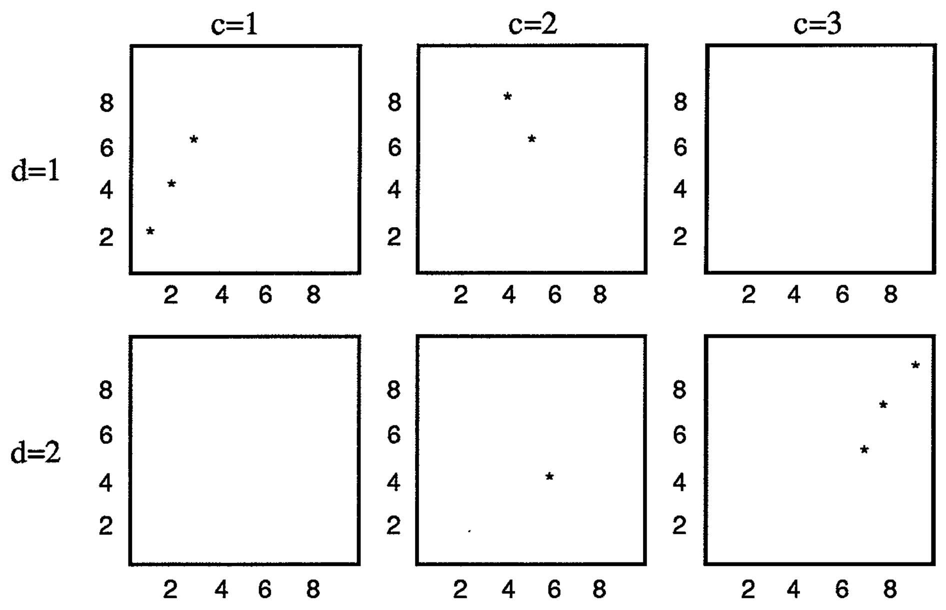

Figure 11. Example of Multiwindow display

The variables **A** and **B** are plotted against each other for each combination of values of **C** and **D**. This kind of plot may reveal third order interaction between the variables. Because of the small data set, Figure 11 isn't quite illuminating. But hopefully it is clear that when many object are involved the multiwindow display might be an interesting tool for rather complex relations.

The figure may be extended by adding the marginal plots of this "table of plots". In fact these marginals are casement displays, summing over **C** and **D** respectively. Compare Figure 10, being the column marginals of Figure 11. Finally notice that the multiwindow display can only be used with four variables simultaneously.

M and N plots

A four dimensional space can also be displayed by 2 and 2 plots (Diaconis and Friedman, 1983). In this method two scatter plots are created and the corresponding objects from the two plots are connected by a line. The length and direction of the lines may show interaction effects. The data from table 1 are plotted this way in Figure 12. It is also possible to make more general M and N plots. For instance Figure 12 could also be displayed by a 3 and 1 plot. The object values for three variables are plotted in a cube-like display and connected with the corresponding values of the fourth variable, that are placed on a straight line. If the values of M and N become larger than two the plot will rapidly get confusing.

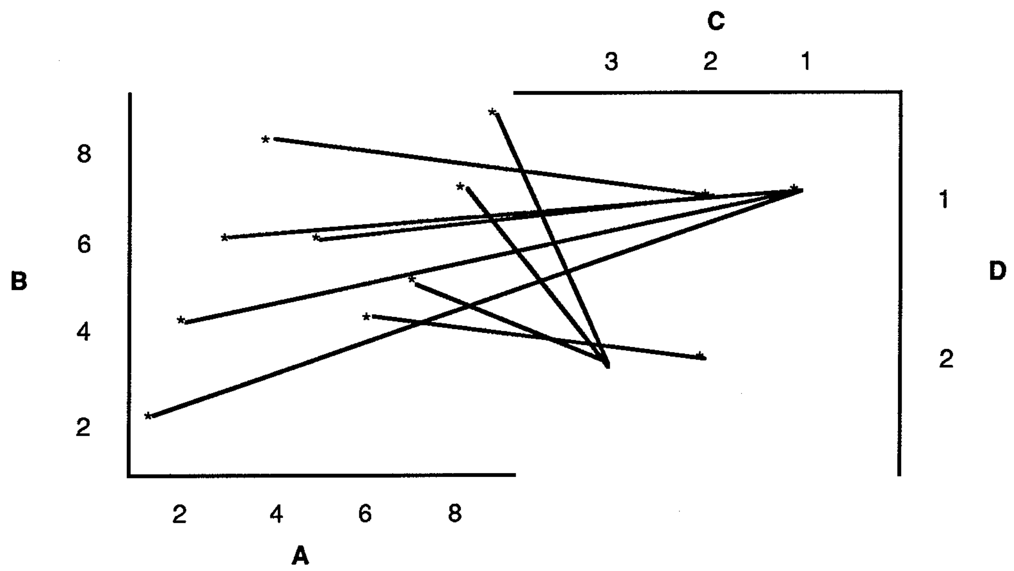

Figure 12. Example of M and N plot

All "symbolic representation" techniques discussed above have in common that they are suitable only with a moderately number of variables or objects. Even if this requirement is met the usefulness of the techniques is restricted to examine the similarity of the objects. An exception is the tree display, which also shows the clustering of variables.

The 'extended scatter plot' approach shows relations between variables but is basically limited to four variables. More variables can be displayed by simply using more multiple windows with all kinds of combinations of variables. This would, however, dramatically increase the difficulty to interpret the overall structure.

When there are many variables it may be better to reduce the dimensionality of the data, before one of the techniques mentioned above is applied.

Dimension reduction

A far more important approach in multivariate analysis is the application of techniques which start by reducing the dimensionality of the data. The *n* object-vectors are projected onto a low-dimensional space. In this low-dimensional space we hope to find all the important aspects of the data (see also Gower & Digby, 1981).

Two widely used techniques for dimension reduction are Principal Components Analysis (PCA) and Factor Analysis (FA). The *m*-dimensional space is reduced to a *p*-dimensional space containing *p* new variables. These variables account for as much variation in the data as possible and are called principal components and factors respectively. Both techniques are interested in the structure of the variables. The most important graphical output therefore is the plot of the component (factor) loadings. In this plot the variables are displayed as vectors projected onto the first *p* principal components (factors). If the length of these vectors becomes nearly 1, the angle between two of them is an indication of the correlation between those variables. Small angles represent large positive correlations. Orthogonal vectors indicate a lack of correlation and vectors pointing in opposite directions have a perfect negative correlation.

A class of techniques that treats objects and variables as equal entities by providing a joint plot with both variables and objects is known by the name of "biplot" (Gabriel, 1971, 1981). The prefix "bi" does not refer to the dimensionality of the low-dimensional space but to the two entities jointly displayed, namely objects and variables. Thus a biplot can also be computed in a three dimensional space (Gabriel & Odoroff, 1986).

To compute a biplot the original datamatrix **Y** is approximated by a matrix **Z** of lower rank (mostly of rank two or three) by means of a singular value decomposition (Eckart & Young, 1936). Each element of this matrix can now be written as an inner product of two vectors with the desired dimensionality: $z_{ij} = a_i' b_j$ ($i=1,...,n$; $j=1,...,m$). The vectors a_i and b_j are displayed in a joint plot and can be considered as 'roweffects' and 'columneffects' respectively.

In Cox and Gabriel (1982) a comparison is made between the biplot and other EDA techniques. They conclude that the biplot is more immediate and more systematic. Patterns seen in the data can sometimes be directly related to specific models.

The data set can also take another form. For instance if the data set only consists of some kind of similarity ratings between the objects, multidimensional scaling techniques (Kruskal & Wish, 1978) are the appropriate tools to use. In this class of techniques the objective is to display the distances between objects in a low-dimensional space. These distances should optimally correspond with the similarity ratings. Thus if the similarity between two objects is large the points representing these objects should be close together in the display.

In an experiment ran by Mezzich and Worthington (1978), 13 subjects were asked to view the same data set displayed by using profiles, stars, faces, Andrews' curves, PCA and multidimensional scaling (MDS). The data set consisted of four groups, each containing 11 objects with measurements on 14 variables. The subjects were asked to uncover this clustering structure (the four groups). The results of this experiment clearly demonstrate that MDS and to a lesser extend PCA were superior to the other techniques in uncovering the clustering structure. Furthermore the subjects indicated that these techniques were easier to use than the other ones.

Although the experiment had several weak points (see Friedman & Rafsky, 1981), it has undoubtedly been demonstrated, that dimension reduction techniques are more suited for the detection of clusters of objects than other techniques.

It has already been mentioned that the aim of multivariate analysis is to obtain a low-dimensional space that represents as much variation in the data as possible. When the correlations between the variables are large, the low-dimensional space may represent more variation of the variables than when these correlations are small. So enlarging the correlations is an important tool in computing the low-dimensional space. To achieve this goal the variables must be transformed. These transformations depend on the measurement level of the variables.

Two techniques that use this idea are HOMALS (Gifi, 1981; Van de Geer, 1985) , an acronym for HOMogeneity analysis by Alternating Least Squares and PRINCALS, acronym for PRINcipal component analysis by Alternating Least Squares (Gifi, 1981, 1983).

HOMALS is developed for data containing only nominal variables. It is actually an extension of correspondence analysis for more than two variables. Contrary to correspondence analysis, however, it provides a plot of objects in the low-dimensional space. Objects that are close together in the plot have similiar profiles, this means that they have chosen approximately the same categories. Furthermore a joint plot of all categories is given. Categories close to each other are approximately chosen by the same objects. A category is actually positioned in the geometrical centre of all objects which have chosen that category.

PRINCALS can be looked upon as ordinary PCA with transformed variables. The program is able to deal with variables with different measurement levels. This means that each variable acquires a transformation dependent on its measurement level. Like PCA a plot is given of the component loadings, showing relations between the variables. Furthermore the objects can be plotted like in HOMALS.

Conclusion

In this paper some static graphical tools are discussed which could be useful within the context of multivariate analysis. First it was seen that methods making use of symbolic representations, although sometimes quite inventive, are not particularly suited to large or even moderately large data sets.

Other graphical tools discussed were scatter plots and some extensions of scatter plots. Scatter plots can handle a rather large number of objects, but with a growing number of variables the number of plots rapidly becomes large, which makes them difficult to interpret. So with many variables techniques that reduce the dimensionality of the data are preferable. The new variables, computed by these techniques, constitute a low-dimensional space and can be portrayed by scatter plot methods.

A final remark. Advanced computer software, like for example SAS Graph and SPSS Graphics, are fully equiped to make all kinds of scatter plots. Especially with color devices one can make clear and comprehensible plots. Although these programs are mainly used for presentation purposes, they can also be useful within the context of multivariate data analysis.

References

Anderson, E. (1960). A semi-graphical method for the analysis of complex problems. Technometrics, 2, 387-392.

Andrews, D.F., (1972). Plots of high-dimensional data. Biometrics, 28, 125-136.

Chambers, J.M. and Kleiner, B. (1982). Graphical Techniques for Multivariate Data and for clustering. In: P.R. Krishnaiah (ed.), Handbook of Statistics Vol. 2. New York: North-Holland.

Chambers, J.M., Cleveland, W.S., Kleiner, B. and Tukey, P.A. (1983). Graphical methods for data analysis. Belmont: Wadsworth International Press.

Chernoff, H. (1973). The use of faces to represent points in k-dimensional space graphically. Journal of the American Statistical Association, 68, 361-368.

Cox, D.R. (1978). Some remarks on the role in statistics of graphical methods. Applied Statistician, 27, 4-9.

Diaconis, P. and Friedman, J.H. (1983). M and N plots. In: M. Rizvi, J. Rustagi and D. Siegmund (eds.). Recent advances in Statistics. New York: Academic Press.

Eckart, C, and Young, G. (1936). The approximation of one matrix by another of lower rank. Psychometrika 1, 211-218.

Friedman, J.H. and Rafsky, L.C. (1981). Graphics for the multivariate two-sample problem. Journal of the American Statistical Association, 76, 277-295.

Gabriel, K.R. (1971). The biplot-graphic display of matrices with applications to principal component analysis. Biometrika, 58, 453-367.

Gabriel, K.R. (1981). Biplot display of multivariate matrices for inspection of data and diagnosis. In: V. Barnett (ed.). Interpreting multivariate data. Chichester: Wiley.

Gabriel, K.R. and Odoroff, C.L. (1986). Some diagnosis of models by 3-D biplots. In: J. de Leeuw, W. Heiser, J. Meulman and F. Critchley (eds.). Multidimensional data analysis. Leiden: DSWO-Press.

Gifi, A. (1981). Non-linear Multivariate Analysis. Leiden: Department of Data Theory.

Gifi, A (1983). PRINCALS user's guide. Leiden: Department of Data Theory.

Gnanadesikan, R. (1977). Methods for Statistical Data Analysis of Multivariate Observations. New York, Wiley.

Gower, J.C. and Digby, P.G.N. (1981). Expressing complex relationships in two dimensions. In: V. Barnett (ed.). Interpreting multivariate data. Chichester: Wiley.

Greenacre, M.J. (1984). Theory and applications of Correspondence Analysis. London: Academic Press.

Hartwig, F. and Dearing, E. (1979). Exploratory Data Analysis. Beverly Hills: Sage University Paper.

Jacob, R.J,K., Egeth, H.E. and Bevan, W. (1976). The face as a data display. Human Factors, 18, 189-200.

Kleiner, B. and Hartigan, J.A. (1981). Representing points in many dimensions by trees and castles. Journal of the American Statistical Association, 76, 260-276.

Kruskal, J.B. and Wish, M. (1978). Multidimensional Scaling. Beverly Hills: Sage University Paper.

Meulman, J. (1986). A distance approach to nonlinear multivariate analysis. Leiden: DSWO-Press.

Mezzich, J.E. and Worthington, D.R.L. (1978). A comparison of graphical methods of multidimensional psychiatric diagnostic data. In: P.C.C. Wang (ed.). Graphical representations of multivariate data. New York: Academic Press.

Ramsay, J.O. (1980). Inside-out displays and more. Presented at symp. Multivariate Data Display, Psychometric Society Meeting, Iowa City.

Tukey, J.W. (1977). Exploratory Data Analysis. Reading, Mass: Addison-Wesley.

Tukey, J.W. and Tukey, P.A. (1983). Some graphics for studying four-dimensional data. Computer Science and Statistics: Proceedings of the 14th Symposium on the Interface. 60-66. New York: Springer-verlag.

Tukey, P.A. (1986). A data analyst's view of statistical plots. Proceedings of the annual meeting of the ASA. Chicago.

Turner, D.W., Seaman, J.M. and Young D.M. (1986). Using polyhedra to graphically display k-dimensional data. Proceedings of the annual meeting of the ASA. Chicago.

Van de Geer, J.P. (1986). Introduction to linear multivariate analysis. Leiden: DSWO-Press.

Van de Geer, J.P. (1985). HOMALS user's guide. Leiden: Department of Data Theory.

Wainer, H. (1979). About faces in factor analysis. BSSR Tech Rep. 547-791.

Wainer, H. (1983). On multivariate display. In: M. Rizvi, J. Rustagi and D. Siegmund, (eds.). Recent advances in Statistics. New York: Academic Press.

Wainer, H. and Thissen, D. (1981). Graphical data analysis. Annual Review of Psychology, 32,191-241.

Wilk, M.B. and Gnanadesikan, R. (1968). Probability plotting methods for the analysis of data. Biometrika, 55,1-17.