ELSEVIER Computational Statistics & Data Analysis 28 (1998) 257–270

COMPUTATIONAL STATISTICS & DATA ANALYSIS

Robust bivariate boxplots and multiple outlier detection

Sergio Zani *, Marco Riani, Aldo Corbellini

Istituto di Statistica, Università di Parma, Via Kennedy 6, 43100 Parma, Italy

Received 1 March 1997; received in revised form 1 February 1998; accepted 2 April 1998

Abstract

In this paper we suggest a simple way of constructing a bivariate boxplot based on convex hull peeling and *B*-spline smoothing. The proposed method shows some advantages with respect to that suggested by Goldberg and Iglewicz (1992). Our approach leads to defining a natural inner region which is completely nonparametric and smooth. Furthermore it retains the correlation in the observations and adapts to differing spread of the data in the different directions. The outer contour, which is based on a multiple of the distance of the inner region from the centre, is robust to the presence of clusters of outliers. We also show how the construction of a bivariate boxplot for each pair of variables can become a very useful tool for the detection of multivariate outliers.

Keywords: Bivariate boxplot; *B*-spline; Convex hull; Multiple outliers; Robust estimate

1. Introduction

The univariate boxplot is universally acknowledged as a very useful tool to summarize univariate data because it provides information concerning location, spread and skewness and also highlights potential outliers (Tukey, 1977). During the last ten years there have been some attempts to construct bivariate contours (Scott, 1985; Press et al., 1986; Becketti and Gould, 1987), but these approaches had serious shortcomings.

Goldberg and Iglewicz (1992) (from now on GI) suggested two main ways of constructing a bivariate boxplot called Relplot and Quelplot. The first is based on a robust elliptic plot. The second uses the so-called quels, which are four separate

* Corresponding author. E-mail: zani@ipruniv.cce.unipr.it.

0167-9473/98/$19.00
PII S0167-9473(98)00040-1

quarter ellipses matched on their major and minor axes in order to ensure continuity and smoothness in the overall graph.

In this paper we suggest an alternative way of constructing a bivariate boxplot which uses the instruments of convex peeling and B-splines.

The structure of the paper is as follows: in Section 2 we describe our proposal for a bivariate boxplot. In Section 3 we compare our method with the existing ones, with special attention to the work of GI. Section 4 is devoted to simulation results. We monitor first the probability of an observation lying outside our outer bivariate contour under the hypothesis of a bivariate normal distribution. Then we analyse the theoretical number of observations left outside the outer contour in relation to the sample size. In Section 5 we apply our method to some well-known data sets containing bivariate outliers, in which traditional methods seem to fail to detect atypical observations. In Section 6 we claim that for p-dimensional data the construction of a bivariate boxplot for each pair of variables can help to find an initial clean data set free from outliers, which can act as a starting point in the forward search techniques (Hadi, 1992; Atkinson, 1994; Riani and Zani, 1996; Atkinson and Riani, 1997) for the detection of multivariate outliers. Final comments are presented in Section 7.

2. Description of the method

In order to construct a bivariate boxplot we must (a) specify an inner region, (b) define a robust centroid and (c) devise a method to construct the outer region. In this section we will examine these three steps in detail.

Step 1: *Definition of the inner region.* The inner region (hinge) is the two-dimensional extension of the interquartile range of the univariate boxplot. In one dimension we take the length of the box which contains 50% of the values. In two dimensions we look for a similar region centered on a robust estimator of location, containing a fixed percentage of the data. A natural and completely nonparametric way of finding a central region in $\mathbb{R}^2$ is through the use of the so-called convex hull peeling (Bebbington, 1978). Barnett (1976) suggested that "the most extreme group of observations in a multivariate sample are those lying on the convex hull (with those on the convex hull of the remaining sample, the second most extreme group, etc.)". The output of the peeling is a series of nested convex polygons (hulls). We call the $(1-\alpha)$%-hull the biggest hull containing not more than $(1-\alpha)$% of the data. (The points on the boundary belong to the hull). Usually, even if the outermost hull assumes very different shapes and is influenced by outliers, the 50%-hull seems to capture the correlation of the two variables. However, if the sample size is small the 50%-hull might not be smooth. In order to overcome this problem we suggest to superimpose around the 50%-hull previously formed a B-spline curve.[1] This leads us to define the inner region as follows:

[1] A B-spline is basically an interpolation cubic curve able to smooth a series of n points given in any order. It uses a four-point control system for each interval of the curve and fulfills the conditions that the first three derivatives are continuous and the distance of the curve from the points is minimal (Ammeraal, 1992).

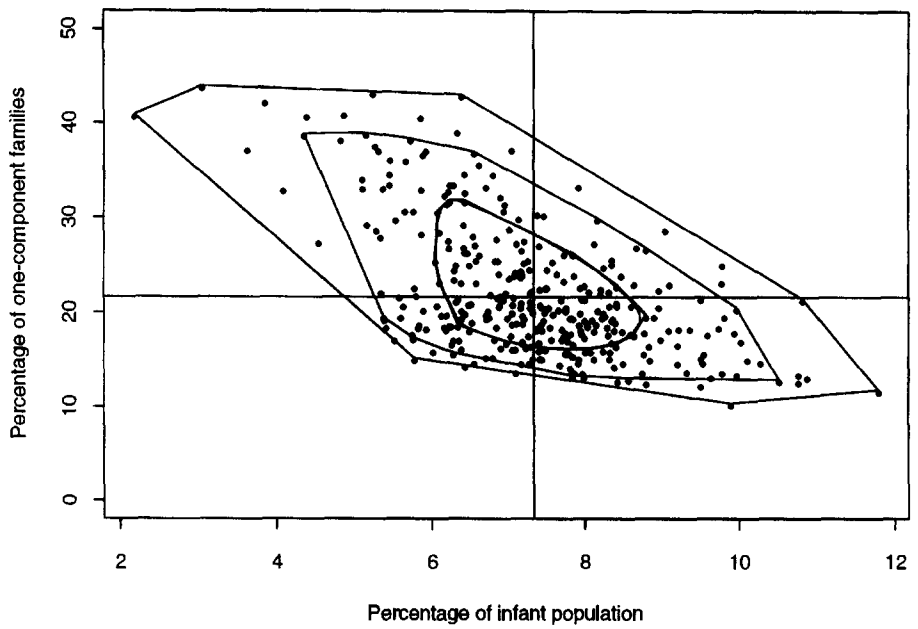

Fig. 1. Plot of the percentage of one-component families versus the percentage of infant population in the 341 municipalities of Emilia-Romagna (Italy) with most extreme, 90% and 50%-hulls. A B-spline curve has been superimposed on the 50%-hull and two straight lines have been drawn through the robust bivariate centroid.

Definition. We call "inner region" the one formed by those units which lie inside or at the boundary of the B-spline curve superimposed on the 50%-hull.

As an illustration of our method we use the data plotted in Fig. 1 (data available on request). On the x axis we have the percentage of infant population ($\leq$10 yr old) of the 341 municipalities of Emilia-Romagna (a region in the north of Italy). On the y axis we have the percentage of one-component families. The bulk of the data, as expected, shows a negative correlation. However, there are many points which seem to depart from an imaginary robust regression line that one could draw through the distribution of the data. Fig. 1 also gives the most extreme hull, the 90%-hull and the 50%-hull. On the 50%-hull we have superimposed a B-spline curve. This inner region is surely free from outliers and robust, but at the same time it keeps the correlation in the data and it allows for different spreads in the various directions.

Step 2: *Definition of the robust centroid.* A robust bivariate centre can be found as the centre of the innermost hull, or from a simple robust criterion such as the minimization of the L_1-norm in $\mathbb{R}^2$ (see Small (1990) for a survey on multidimensional medians) or through the intersection of two robust regression lines (for example least median of squares lines (Rousseeuw and Leroy, 1987)). Another viable possibility is the use of Tukey's halfspace median, which is the point with largest halfspace depth (Tukey, 1975). For computational details, see Rousseeuw and Ruts (1996, 1997).

Our suggestion is to find a robust bivariate centroid using the arithmetic means of the observations inside the inner region. In this way we exploit both the effi-

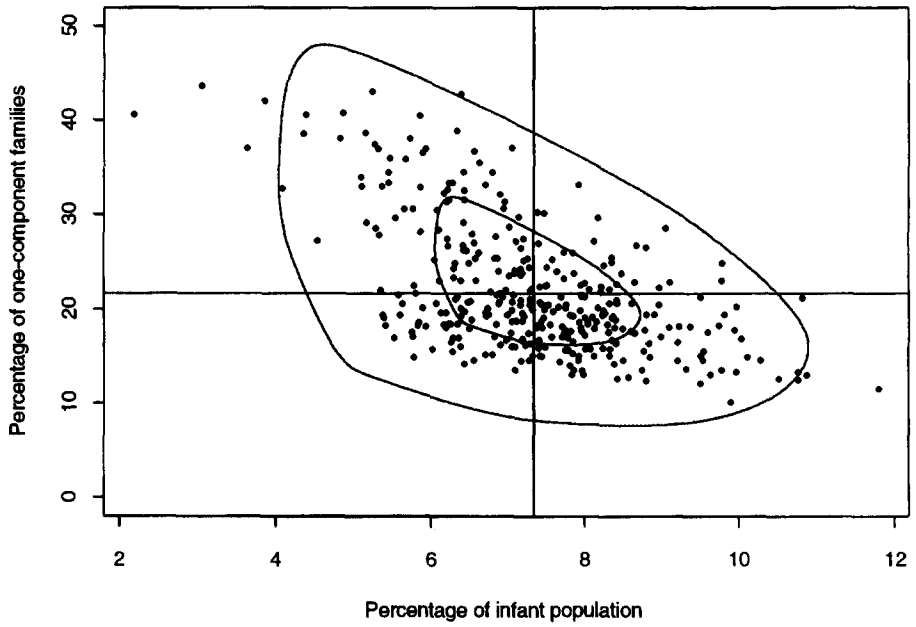

Fig. 2. Bivariate boxplot of the data reported in Fig. 1.

ciency properties of the arithmetic mean and the natural trimming offered by the hulls. In Fig. 1 we have also drawn the two straight lines parallel to the Cartesian axes which intersect this bivariate robust centroid. Another viable alternative consists in using the intersection of the two least squares lines built with the observations forming the inner region. This latter approach, which has the advantage of taking into account the relation between the two variables, will be explored in subsequent work.

Step 3: *Definition of the outer region.* Once we have found a robust bivariate centre and a curve containing half of the data (hinge) we must devise a criterion in order to build an outer contour which discriminates between "good" and "bad" observations.

In the traditional univariate boxplot we define as outliers those observations which lie outside the following interval: $[x_{0.25} - 1.5 \times IR, x_{0.75} + 1.5 \times IR]$ where IR is the interquartile range ($IR = x_{0.75} - x_{0.25}$). As is well-known, if the data come from the normal distribution, the former interval contains 99.3% of the values.

In two dimensions, in order to allow a degree of asymmetry, we suggest to build a bivariate outer contour using a multiple of the distance of the hinge from the robust centre. The purpose is to find a contour which leaves outside the (small) proportion of the data which can be considered as atypical. Furthermore, under the hypothesis of bivariate normality, this contour can be interpreted as a probability contour at $(1-\alpha)$ level with very small α (say close to 0.01). Lastly this contour is robust to departures from normality. In the appendix we show that in a bivariate normal distribution, in order to find an outer contour which leaves outside a percentage of observations close to 1%, we must multiply the distance of the hinge to the robust centre by

1.58. This coefficient (which we call l) is independent of the correlation parameter ρ of the bivariate normal distribution (see the Appendix). In one dimension this would correspond to the separate use of the distances between the upper or lower quartile and the median in order to compute the upper or lower truncation points (Lanska and Kryscio, 1996). Using this coefficient we obtained the outer region plotted in Fig. 2.

The graph clearly shows that the shape of the outer region (which is completely nonparametric and smooth) adapts to the differing spread of the data in the different directions.

In this section we have constructed a bivariate boxplot starting from an inner region containing not more than 50% of the data. Of course nothing prevents us from using an inner region based on a larger proportion of the data (say 75%). In this case the method follows the same lines with the only difference that we must use a different coefficient l to build the outer contour (see the Appendix).

The program to build our bivariate boxplot was developed under S-Plus version 3.2. Aware of the inefficiency of S-Plus in providing fast-loop algorithms, all functions which require heavy looping or multiple nested cycles, like the computation of B-spline coefficients, were developed under pure C. While many simulations on large datasets could be time consuming (i.e.: 1000 simulations with 1000 observations each, require a little less than an hour on an Intel-Pentium 150 (R)-based machine with 32 MB Ram) computing hulls, B-splines and outliers on a single dataset of 1000 observations is nearly instantaneous (2–3 s). Upon request the authors will provide the C and S-plus source code.

3. Comparison with existing methods

In order to find a robust inner region Rousseeuw (1984) suggested to find the ellipse with the smallest area (or more generally with p-dimensional data the minimum volume ellipsoid) containing half of the observations. However, the ellipse does not allow differing spread in the different directions of the data. For example as emerges clearly from Fig. 1, the spread of the data from the robust centroid is not symmetric. Consequently a robust confidence ellipse does not seem to be appropriate.

GI, in order to allow for different variability in various directions, use 4 quarter ellipses (quels) matched on the minor and major axes. This procedure, which has the purpose of giving an additional degree of asymmetry, is based on bivariate biweight estimators and is computationally cumbersome.

More precisely GI suggest two methods for the construction of the bivariate boxplot: the first is based on a robust elliptic plot (Relplot) the second on 4 quarter separate ellipses (Quelplot). To draw a Relplot one needs two robust estimators of location T_x and T_y; the ratio between the two estimators of scale S_x, S_y, one estimator of the correlation R and a constant which regulates the distance of the fence from the hinge.

To draw a Quelplot one needs two additional parameters P_1 and P_2 which are the proportions of the total standard deviation due to the residuals in the positive

direction of the major and minor axes of the ellipse. Globally, therefore, 6 quantities must be estimated.

In our method we simply need a robust estimate of the bivariate centroid. The contour of our hinge is (a) robust, (b) completely nonparametric, being constructed by a spline superimposed on the 50% hull which does not require the estimation of parameters, (c) has a sufficient degree of smoothness and (d) is not necessarily symmetric.

GI in order to draw the fence use the "single nonoutlying point farthest from the centre". This means that we can always find one observation which lies on the threshold which declares one unit as outlier. Therefore, with small sample sizes the confidence region can vary considerably if we move the coordinates of a point which lies on the fence.

4. Simulation results

The B-spline which is superimposed on the 50%-hull always lies inside it. Therefore the theoretical coefficient we find is based on a region that may contain a little less than 50% of the data. This means that we expect that in a bivariate normal distribution the percentage of data which lie outside the outer contour is a little bit greater than 1% when we use $l = 1.58$. Table 1 reports the results of 5 simulations for the following 5 values of ρ: $\rho = 0.1, 0.3, 0.5, 0.7, 0.9$. Of course, we can suppose without loss of generality that the standard deviations of the two variables are equal to 1 and the true means are equal to 0, because convex hull peeling is invariant under linear transformations of the data (Donoho and Gasko, 1982; Ruts and Rousseeuw, 1996). From the results of Table 1 it emerges that when the sample size is equal to 1000 the average Monte-Carlo percentage of observations which lie outside of the outer region is around 1.29%.

Another issue to examine is the percentage of observations outside the outer contour (outside rate per observation) for different values of the sample size (n). The sample sizes we considered are: 50, 100, 200, 400, 500, 800 and 1000. As is well-known (Hoaglin et al., 1986; Goldberg and Iglewicz, 1992), the outside rate per observation is a decreasing function of n. From the results of Fig. 3, which reports the outside rate per observation for different values of n, it emerges that we experience the same decreasing pattern. The coefficient (l) which asymptotically enables

Table 1
Monte Carlo average number of observations ($\bar{m}$) outside the outer contour using $l = 1.58$ (1000 simulations of 1000 units each)

ρ	$\bar{m}$
0.1	12.892
0.3	13.004
0.5	12.763
0.7	12.681
0.9	12.428

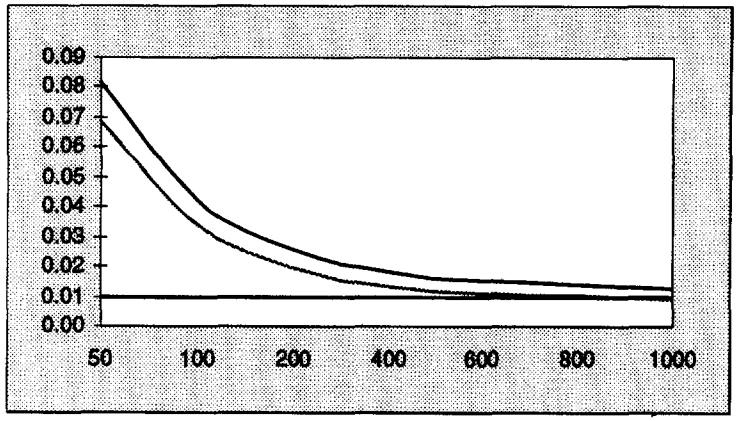

Fig. 3. Outside rate per observation (in percent) versus sample size. Simulated estimates from bivariate Gaussian samples with $\rho = 0.5$. The upper line is referred to the 1% theoretical coefficient ($l = 1.58$), the lower line to $l = 1.68$ (empirical coefficient).

us to have an outside rate per observation of 1% is 1.68. In the examples of the next section we use this value to draw the fence.

The coefficients corresponding to asymptotic outside rates per observation of 10% and 5%, are 0.86 and 1.13, which are a little bit bigger than the theoretical coefficients of 0.82 and 1.08 (see the Appendix).

5. Examples

Our method is directed in a particular way towards big data sets. When the sample size is very small ($n \leq 20$) it might happen that the 50%-hull contains far fewer points than 50% of the data points, and hence no longer represents the shape of the bulk of the data. In this section however, we show that our bivariate boxplot seems to work well also for small sample sizes. The first bivariate data we consider concerns Log light intensity versus Log effective surface temperature for the 47 stars of the star cluster CYG OB1 (Rousseeuw and Leroy, 1987, pp. 27–28). This example has also been used by GI because in this case an elliptic plot based on least squares estimators distorts the data structure. GI identify five stars as clear outliers. In their Relplot, observation 14 lies a little outside the fence and unit 9 is used to build the outer contour. Therefore it is forced to lie on it. In their Quelplot, observation 14 lies on the boundary and unit 9 falls inside the outer contour. Fig. 4 reports our bivariate boxplot and shows that even when the sample size is small our method clearly enables us to identify correctly the structure of the data. Similar to GI, units 34, 30, 20, 11 and 7 are clearly recognized as outliers. In our approach no unit is forced to lie on the outer contour and, while observation 14 falls outside, unit 9 remains inside.

The second data set we consider is known in the literature as the Brain and Body weight data and concerns the Log of the weights of body and brain for 28 species

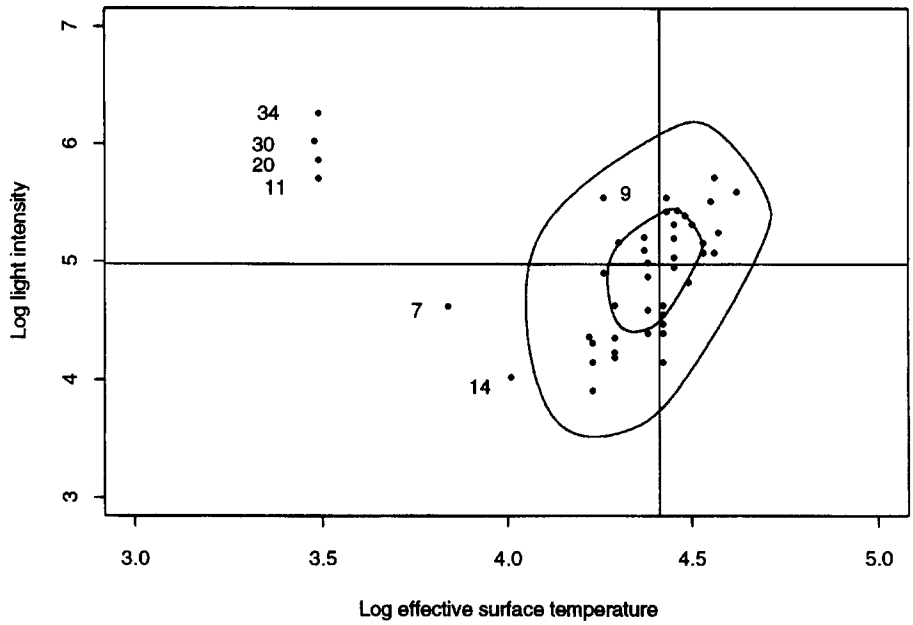

Fig. 4. Bivariate boxplot of Log light intensity and Log effective surface temperature for the 47 stars of the star cluster CYG OB1.

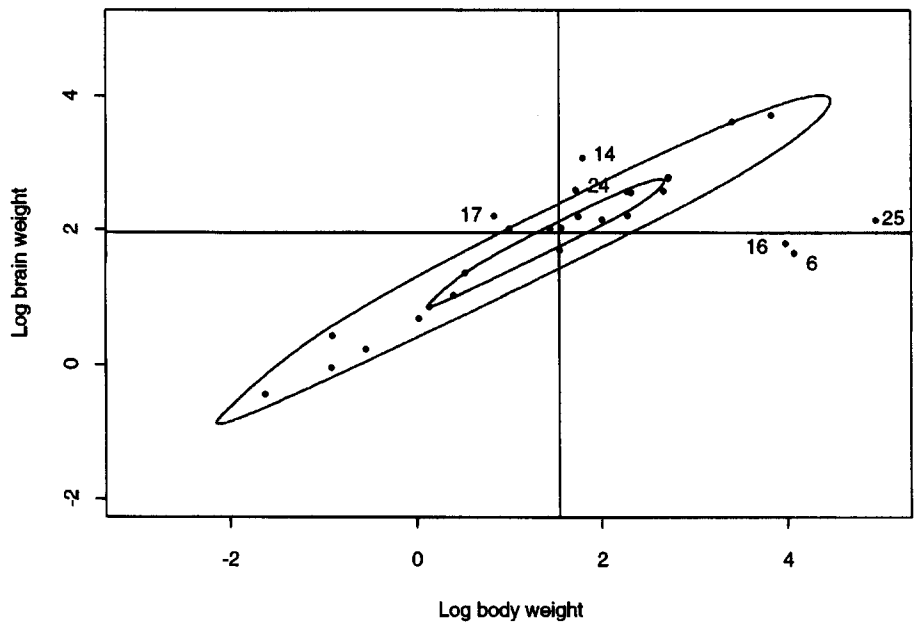

Fig. 5. Bivariate boxplot of Log brain weight and Log body weight (Observations detected as outliers are identified with their number).

(the raw data can be found in Rousseeuw and Leroy (1987, p. 58)). Fig. 5 shows that the majority of the data follow a clear pattern with some exceptions on the lower right region and on the upper part of the plot. Rousseeuw and Van Zomeren (1990) build the minimum volume ellipsoid and detect observations 6, 16, 25, 14 and 17 as outliers. Observation 24 lies on the boundary of their 97.5% tolerance ellipse. As is clearly shown in Fig. 5, our outer region enables us to detect the 5 atypical observations formerly identified. A difference from Rousseeuw and Van Zomeren (1990) is that our bivariate boxplot observation 24 lies a little outside the outer region.

Remark. It follows from the analysis of the outside rate per observation that, when the sample size is small, we expect that the percentage of observations which lie outside the outer region is larger than 1%. However, in order to be conservative, we always suggest using the coefficient 1.68.

6. Bivariate and multivariate outliers

The focus of this paper concerns bivariate data and the detection of bivariate outliers. However the analysis of atypical bivariate observations can offer us a good starting point for the detection of multivariate outliers. The modern procedure of detecting atypical p-dimensional observations is based on the following two steps (e.g. Hadi, 1992, 1994):

(1) definition of a clean data set (initial or basic subset) free from outliers of dimension $p+1$;

(2) iterative inclusion of units in the clean data set until a stopping rule is satisfied.

One problem is the choice of the clean data set with which to start the forward iterative inclusion. For example Atkinson (1994), in order to avoid this choice, suggested to extract a subset of observations at random and to repeat the forward search several times.

With p dimensional data the construction of bivariate boxplots for each couple of variables enables us to define as "bivariate clean data sets" the subsets containing the units inside the $1-\alpha$ outer contours. In this paper we have constructed a bivariate boxplot with an outer region which, in the presence of bivariate normality, leaves outside a percentage of observations close to 1%. This threshold however can conveniently be decreased in order to increase the probability that the initial clean data set is free from outliers. In large samples we suggest to use 90% outer regions in each scatter plot. In small samples 75% outer regions might be a proper choice.

A natural definition of the multivariate initial subset (or basic subset) is the following (Riani and Zani, 1996):

Definition. We call the initial subset of multivariate clean observations the one formed by the intersection of the subsets of bivariate clean data in each of the $p(p-1)/2$ pairs of variables.

The observations which lie at least once outside the outer contour can be considered as potential outliers and can be removed from the basic subset.

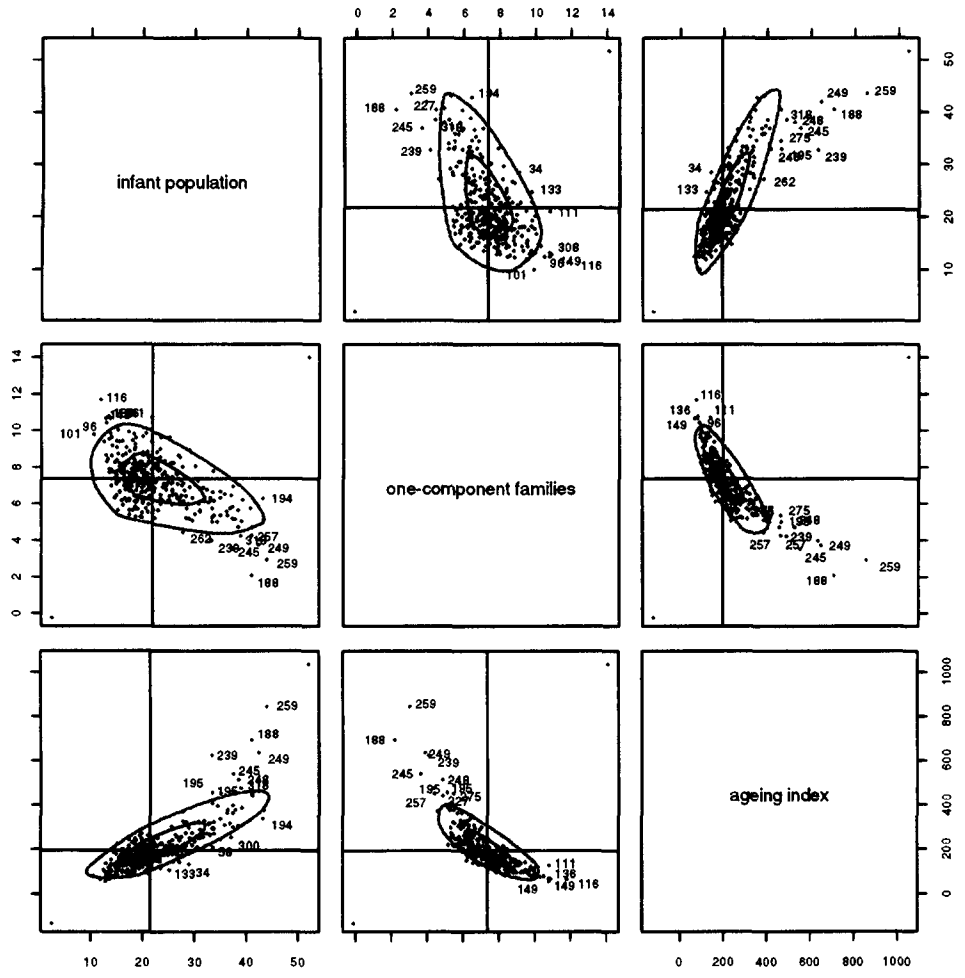

Fig. 6. Scatter plot matrix of three demographic variables and bivariate boxplots with 90% outer contours.

Usually outliers form a small part of the overall sample. Therefore it is not sensible to start the forward search with a clean data set of dimension $p+1$ especially when the sample size is large. The former criterion enables us to start with a basic subset of a certain dimension. This provides computational savings and simplifications in the analysis of potential outliers.

Of course one observation which is a multivariate outlier might always lie at the boundary of the bivariate confidence regions without ever falling outside. In the rare event that the basic subset contained some atypical observations they could be immediately detected with a forward search in which units can leave the basic subset (Hadi and Simonoff, 1993; Atkinson, 1994; Atkinson and Riani, 1997).

As an illustration of the suggested method for the detection of multivariate outliers, let us suppose that the data reported in Fig. 1 belong to a three-dimensional data set in which the third variable is an ageing index. Fig. 6 shows the bivariate boxplots

for each pair of variables. In order to be conservative we have used the coefficient 0.82 which has a theoretical outside rate per observation equal to 0.10. All the units marked with a number are excluded from the initial clean data set. Initial Mahalanobis distances are calculated using a centroid based on the statistical units forming the basic subset. Observations belonging to the group of potential outliers are iteratively included in the initial subset. Results can be displayed through simple graphs of a variety of statistics monitored along the forward search algorithm (Riani and Zani, 1996; Atkinson and Riani, 1997).

7. Conclusions

In this paper we have suggested a simple and transparent method to construct a bivariate boxplot based on convex hull peeling and B-spline smoothing. These tools lead us to define a natural inner region (hinge) which is completely nonparametric and smooth. Furthermore it keeps the correlation in the observations and it adapts to the different spread of the data in the different directions. The outer region (fence), which is based on a multiple of the distance of the points lying on the hinge to the centre, is robust to the presence of clusters of outliers. In addition, under the hypothesis of bivariate normality, the fence can be interpreted as a bivariate confidence contour at the 99% level.

The proposed method has been applied with success to some well-known data sets formerly analysed in the literature concerning atypical values.

The focus of the paper is mainly directed to bivariate outliers. For p-dimensional data, however, the construction of a bivariate boxplot for each pair of variables can be a very useful tool for defining an initial clean data set, which can then act as a starting point in the usual forward search techniques for the detection of multivariate outliers.

Appendix

Let O be the origin of the cartesian axes and $\overline{OA} = k_{0.5}$ be the distance of a point A from the centre O which lies on the ellipse of a bivariate standard normal distribution associated with a constant probability density contour of 0.50. Let $\overline{OB} = k_{0.99}$ be the distance of a point B from O which lies on the ellipse associated with a probability density contour of 0.99. Segment $\overline{OB}$ is such that it intersects point A (Fig. 7).

In this appendix we prove that the probability that an observation lies inside the ellipse which intersects the point which has a distance from the centre equal to $\overline{OB} = \overline{OA} + l\overline{OA}$ is approximately equal to 0.99 when $l = 1.58$.

Suppose initially that $\rho = 0$. In order to compute the radius $\overline{OA}$ of the circle which contains 50% of the values we must solve the following bivariate integral:

$$\int\int_{\{z_1,z_2:\ z_1^2+z_2^2\le\overline{OA}^2\}} \frac{1}{2\pi}\exp\{-\tfrac{1}{2}(z_1^2+z_2^2)\}\,dz_1\,dz_2 = 0.5.$$

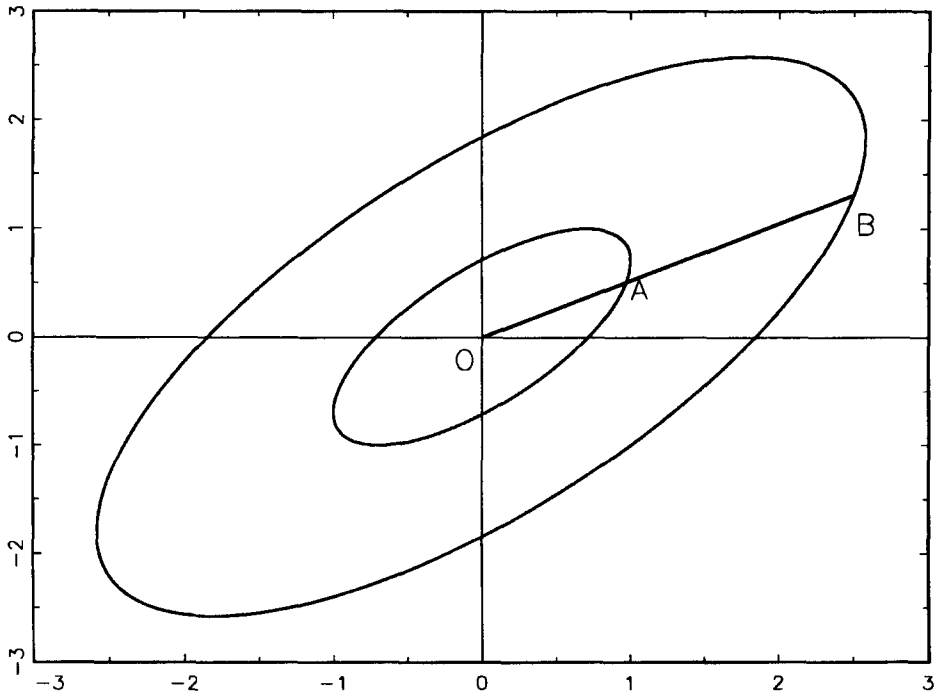

Fig. 7. Ellipses associated with two density probability contours in a bivariate standard normal distribution.

Passing to polar coordinates we have

$$\int_0^{\overline{OA}} r \exp\{-\tfrac{1}{2}r^2\}\, dr = 0.5.$$

Notice that after the change of variables $r^2 = x$ the expression under the integral sign can be immediately recognized as χ^2 with 2 degrees of freedom. We find $\overline{OA} = 1.177$.

If we want the radius $\overline{OB}$ corresponding to 0.99 we easily obtain that $\overline{OB} = 3.03485$.

If $\rho \neq 0$ it is easy to prove that the quadratic form in the exponent of the bivariate normal distribution (x_1, x_2) after a rotation in which the new coordinates of the cartesian axes pass through the axes of the ellipse, can be written as follows:

$$\frac{x_1^2}{1+\rho} + \frac{x_2^2}{1-\rho} = k_{1-\alpha}^2$$

where $k_{1-\alpha}^2$ is the coefficient associated with the ellipse containing a fraction $(1-\alpha)$ of the values. The distance of a point which lies on this ellipse from the centre O which intersects the straight line $(y = mx)$ where m denotes a generic slope is:

$$\sqrt{\frac{k_{1-\alpha}^2(1-\rho^2)(1+m^2)}{(1-\rho)+m^2(1+\rho)}}. \tag{A.1}$$

Table 2
Theoretical (l_t) versus empirical (l_e) coefficients associated with different probability contours in a standard normal distribution

Prob.	l_t	l_e
0.99	1.58	1.68
0.95	1.08	1.13
0.90	0.82	0.86
0.75	0.41	0.43

The coefficient l which enables one to pass from the hinge to the outer contour must satisfy the following equation:

$$l = (\overline{OB} - \overline{OA})/\overline{OA}.$$

Given that from every point ρ and m are fixed, from Eq. (A.1) we have

$$l = (k_2 - k_1)/k_1. \tag{A.2}$$

This means that even if the distance $\overline{OA}$ depends on ρ and from the position of the point on the ellipse the coefficient l is independent of both. Consequently we can argue supposing $\rho = 0$. Substituting in Eq. (A.2) $k_2 = 3.03485$ and $k_1 = 1.177$ we find $l = 1.58$.

In Table 2 we give the theoretical and empirical coefficients which enable us to pass from our inner region defined in Section 2, to an outer contour containing asymptotically 75%, 90%, 95% or 99% of the observations in a bivariate normal distribution.

Acknowledgements

This research was partially supported by a grant from the Italian Ministry of University (MURST). The authors wish to thank A.C. Atkinson for helpful discussion on an earlier draft of this paper and a referee for helpful comments.

References

Ammeraal, L., 1992. Programming Principles in Computer Graphics. Wiley, New York.

Atkinson, A.C., 1994. Very fast robust methods for the detection of multiple outliers. JASA 89, 1329–1339.

Atkinson, A.C., Riani, M., 1997. Bivariate boxplots, multiple outliers, multivariate transformations and discriminant analysis: the 1997 Hunter lecture. Environmetrics 8, 583–602.

Barnett, V., 1976. The ordering of multivariate data. (with discussion) JRSS, Ser. A 139, 318–339.

Bebbington, A.C., 1978. A method of bivariate trimming for robust estimation of the correlation coefficient. Appl. Statist. 27, 221–226.

Becketti, S., Gould, W., 1987. Rangefinder box plots. A note. Amer. Statist. 41, 149.

Donoho, D.L., Gasko, M., 1992. Breakdown properties of location estimates based on halfspace depth and projected outlyingness. Ann. Statist. 20, 1803–1827.

Goldberg, K.M., Iglewicz, B., 1992. Bivariate extensions of the boxplot. Technometrics 34, 307–320.
Hadi, A.S., 1992. Identifying multiple outliers in multivariate data. JRSS, B 54, 761–771.
Hadi, A.S., 1994. A modification of a method for the detection of outliers in multivariate samples. JRSS, B 56, 393–396.
Hadi, A.S., Simonoff, J.S., 1993. Procedures for the identification of multiple outliers in linear models. JASA 88, 1264–1272.
Hoaglin, D.C., Iglewicz, B., Tukey, J.W., 1986. Performance of some resistant rules for outlier labeling. JASA 81, 991–999.
Lanska, D., Kryscio, R.J., 1996. Modified box plots for asymmetric distributions. Working paper, University of Kentucky.
Press, W.H., Flannery, B.P., Teukolsky, S.A., 1986. Numerical Recipes: The Art of Scientific Computing. Cambridge University Press, Cambridge, MA, pp. 532–534.
Riani, M., Zani, S., 1996. An iterative method for the detection of multiple outliers. Proc. Conf. Analysis of Multidimensional Data, Naples, Metron (forthcoming).
Rousseeuw, P.J., 1984. Least median of squares regression. JASA 79, 871–880.
Rousseeuw, P.J., Leroy, A., 1987. Robust Regression and Outlier Detection. Wiley, New York.
Rousseeuw, P.J., Ruts, I., 1996. Bivariate location depth. Appl. Statist. (JRSS-C) 45, 153–168.
Rousseeuw, P.J., Ruts, I., 1997. Constructing the bivariate Tukey median. Statistica Sinica, Accepted.
Rousseeuw, P.J., van Zomeren, B.C., 1990. Unmasking multivariate outliers and leverage points. (with discussion) JASA 85, 633–651.
Ruts, I., Rousseeuw, P.J., 1996. Computing depth contours of bivariate point clouds. Comput. Statist. Data Anal. 23, 153–168.
Scott, D.W., 1985. Frequency polygons: theory and applications. JASA 80, 348–354.
Small, C.G., 1990. A survey of multidimensional medians. Internat. Statist. Rev. 58, 263–277.
Tukey, J.W., 1975. Mathematics and the picturing of data. Proc. Internat. Congress Math. Vancouver 2, 523–531.
Tukey, J.W., 1977. Exploratory Data Analysis. Addison-Wesley, Reading, MA.