UNIVERSITY OF BIRMINGHAM

University of Birmingham Research at Birmingham

# On a scale-scale plot for comparing multivariate distributions

Guha, Pritha; Chakraborty, Biman

*DOI:*
10.1080/03610926.2020.1759639

*License:*
Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

*Document Version*
Peer reviewed version

*Citation for published version (Harvard):*
Guha, P & Chakraborty, B 2022, 'On a scale-scale plot for comparing multivariate distributions', *Communications in Statistics - Theory and Methods*, vol. 51, no. 5, pp. 1270-1289.
https://doi.org/10.1080/03610926.2020.1759639

Link to publication on Research at Birmingham portal

**Publisher Rights Statement:**
This is an Accepted Manuscript version of the following article, accepted for publication in Communications in Statistics - Theory and Methods. Pritha Guha & Biman Chakraborty (2022) On a scale-scale plot for comparing multivariate distributions, Communications in Statistics - Theory and Methods, 51:5, 1270-1289, DOI: 10.1080/03610926.2020.1759639. It is deposited under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

**General rights**
Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

•Users may freely distribute the URL that is used to identify this publication.
•Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
•User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
•Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

**Take down policy**
While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Download date: 07. May. 2024

# ON A SCALE-SCALE PLOT FOR COMPARING MULTIVARIATE DISTRIBUTIONS

Pritha Guha[1] and Biman Chakraborty[2]
[1] Productions, Operations and Decision Sciences
XLRI, Xavier School of Management,
C.H. Area(East), Jamshedpur - 831001
Jharkhand, India
email: pritha@xlri.ac.in
[2]School of Mathematics, University of Birmingham, Edgbaston,
Birmingham B15 2TT, United Kingdom
email: B.Chakraborty@bham.ac.uk

ABSTRACT

In this paper, we propose a scale-scale plot to compare multivariate distributions. These scale-scale plots can be viewed as a multivariate analogue of quantile-quantile plots and we illustrate their use as a visualisation tool to validate distributional assumptions for multivariate data as well as to compare the distributions of two multivariate samples. We discuss some characterisations of the proposed plots under elliptically symmetric distributions and based on those results, some visual tests of location and scale are proposed as further applications of these scale-scale plots. For the test of location problem, we present a small study of power using simulations.

Key Words: Central rank regions, elliptically symmetric distributions, quantile quantile plots, tests of location, tests of scale

**ON A SCALE-SCALE PLOT FOR COMPARING MULTIVARIATE DISTRIBUTIONS**

Pritha Guha[1] and Biman Chakraborty[2]
[1] Productions, Operations and Decision Sciences
XLRI, Xavier School of Management,
C.H. Area(East), Jamshedpur - 831001
Jharkhand, India
email: pritha@xlri.ac.in
[2]School of Mathematics, University of Birmingham, Edgbaston,
Birmingham B15 2TT, United Kingdom
email: B.Chakraborty@bham.ac.uk

ABSTRACT

In this paper, we propose a scale-scale plot to compare multivariate distributions. These scale-scale plots can be viewed as a multivariate analogue of quantile-quantile plots and we illustrate their use as a visualisation tool to validate distributional assumptions for multivariate data as well as to compare the distributions of two multivariate samples. We discuss some characterisations of the proposed plots under elliptically symmetric distributions and based on those results, some visual tests of location and scale are proposed as further applications of these scale-scale plots. For the test of location problem, we present a small study of power using simulations.

Key Words: Central rank regions, elliptically symmetric distributions, quantile quantile plots, tests of location, tests of scale

# 1 Introduction

One of the widely used graphical methods for comparing univariate distributions is the quantile-quantile (Q-Q) plot that matches the quantiles of one distribution with the same quantiles of the other. The Q-Q plots, which were proposed by Wilk and Gnanadesikan (1968), are quite useful in revealing location and scale differences as well as identifying outliers. Though there is an extensive literature on univariate Q-Q plots, (see Barnett (1976), Cook and Weisberg (1982), Cleveland (1993), Marden (2004), for some detailed discussions and examples), there are very few proposed generalisations to the multivariate distributions. Most of the multivariate procedures are based on dimension reduction techniques and are used to compare some specific reference distributions, such as the multivariate normal distribution and depend on their properties.

Healy (1968) used squared Mahalanobis distances of the observations from the sample mean vector to assess multivariate normality. These squared distances are approximately distributed as chi-squared random variables and a Q-Q plot can be constructed to assess that. Andrews, Gnanadesikan, and Warner (1973) proposed Q-Q plots based on the directions and the magnitudes of the observations. The magnitudes or equivalently the squared distances are approximately distributed as chi-squared random variables as before and the angles obtained from the direction vectors are uniformly distributed. There are some other multivariate Q-Q plotting techniques available in the literature, which are based on assessing the commonality of the shape of the marginal distributions for certain commonly used multivariate distributions. For a detailed discussion on graphical methods for assessing multivariate distributional shape, see chapters 5 and 6 of Gnanadesikan (1977).

In a completely different approach, Easton and McCulloch (1990) proposed a generalisation of multivariate Q-Q plots based on matching the data with simulated observations from the reference distribution. For a $d$-dimensional data set $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$, the procedure is to find a permutation $\sigma^*$ of $\{1, 2, \cdots, n\}$ a $d \times d$ matrix $\mathbf{A}$ and a $d \times 1$ vector $\boldsymbol{b}$ that solves

$$\min_{\mathbf{A}, \boldsymbol{b}, \sigma} \sum_{i=1}^{n} ||\boldsymbol{Y}_i - \mathbf{A}\boldsymbol{X}_{\sigma(i)} - \boldsymbol{b}||^2$$

where $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n$ is a random sample from the reference distribution. Suppose the $\boldsymbol{X}_i^*$'s are the best matching of the data set to the reference sample where $\boldsymbol{X}_i^* = \mathbf{A}^*\boldsymbol{X}_{\sigma^*(i)} + \boldsymbol{b}^*$. They suggested to display the matched pairs $(\boldsymbol{X}_i^*, \boldsymbol{Y}_i)$ using either coordinatewise Q-Q plots or some distance based Q-Q plots. One of the main problem in this approach is that it cannot be used to compare two multivariate samples. Visualising co-ordinatewise Q-Q plots can be difficult if the dimension $d$ is large. However the proposed method is invariant under affine transformations and can be quite useful in picking up the deviations in shape from the reference sample when the dimension is not very large. There are some more graphical tools available in the literature to detect non-normality for high dimensional data; see, for example, Liang et al. (2004), Fang et al. (1998), Liang and Ng (2009). However all of these methods heavily depend on different characterisations of multivariate normal distribution and cannot be extended to a large class of multivariate distributions.

Singh, Tyler, Zhang and Mukherjee (2009) proposed a notion of quantile scale curves to perform visual tests. Again consider the $d$-dimensional random sample $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ and denote by $\triangle_i$ the volume of the simplex formed by $(d+1)$ points $\boldsymbol{X}_{i_1}, \cdots, \boldsymbol{X}_{i_{d+1}}$ with $\alpha_i = \{i_1, \cdots, i_{d+1}\}$, a $(d+1)$ subset of indices of $\{1, \cdots, n\}$, for $i = 1, \cdots, N$ and $N = \binom{n}{d+1}$. Then their quantile scale curve is defined as $qsc(t) = \triangle_{([Nt])}$ for $0 \leq t \leq 1$. They have used the quantile scale curve to detect linear and non-linear association between two groups of variables graphically. They have also described quantile scale curve based testing plans for location shift problem in a symmetric population set up and multivariate scale comparison problem for two given multivariate samples. However, this proposal is not a generalisation of Q-Q plots as it does not compare the multivariate distributions.

In an approach based on the spatial quantiles, Marden (1998) proposed a bivariate Q-Q plot by drawing arrows from (bivariate) normal quantiles corresponding to bivariate ranks of the observations to the actual values of the corresponding observations as a check for normality. Chakraborty (2001) pointed out that Marden's plots are not affine invariant and hence might lead to erroneous inference for highly correlated data. He suggested a modification based on a transformation retransformation (TR) approach to construct an affine invariant Q-Q plot for bivariate data. However, these suggestions cannot be extended to higher dimensions in any natural way. In a recent work Dhar, Chakraborty and Chaudhuri (2014) proposed a method to construct a Q-Q plot for a $d$-dimensional multivariate dataset as a collection of $d$ two dimensional plots, each plot corresponding to a component of the multivariate empirical spatial quantile of the data. They also proposed a test based on the norms of the spatial quantiles for comparing multivariate distributions. Their proposed test statistic is related to the arrow lengths of the arrow plots proposed by Marden (1998) described earlier. In a different development Balanda and MacGillivray (1990) proposed spread-spread plots to compare univariate distributions where the spread functionals preserve the spread ordering of Bickel and Lehman (1979). The spread-spread plots are quite useful in detecting the changes in the shape of the distributions at the peak or at the tails and it displays a growth pattern. Wang and Zhou (2012) have considered a general depth function $D(x, F)$ and its corresponding central region $C_{F,D}(p)$ for a distribution $F$ and with any measure $m(\cdot)$ in $\mathbb{R}^d$, they have considered $\lambda_F(p) = m(C_{F,D}(p))$. They have also shown that, $\lambda_F(p)$ is a generalised quantile function. In their paper through Theorem 2.1 and

2.2, they have shown that $\lambda_F(p)$ is affine invariant, which would be a desirable property for any measure which has been developed to describe or compare multivariate distributions. Discussion on desirable properties of data depth functions is present in Zuo and Serfling (2000). Our results in Theorem 2.1 and Theorem 3.1 may appear similar, but our results are with respect to our definition of central rank regions and volume functionals. As the definitions of our central rank regions are based on affine invariant spatial ranks, the approach is different in nature.

In this paper, recent developments in multivariate quantiles and ranks and the concept of univariate spread-spread plots are combined to propose graphical methods of comparing multivariate distributions. In Section 2, we define a notion of multivariate rank vector and a scale curve following the idea of Liu et al. (1999) and study its properties under elliptically symmetric distributions. In Section 3, scale-scale plots are proposed and we discuss their uses with examples in the one sample and two sample problems. These scale-scale plots are also quite useful in detecting the deviations in regard to peakedness or tail behaviour. In Section 4, we propose some visual tests of multivariate location and scale under elliptic symmetry of the distributions as some further applications of our scale-scale plots. All proofs are relegated to the appendix.

# 2 Multivariate Rank Regions:Definitions and Some Basic Properties

We begin with a vector-valued notion of multivariate sign by generalising the univariate sign function $sign(x) = x/|x|$, for $x \neq 0$ with the definition

$$Sign(\boldsymbol{x}) = \begin{cases} \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} & \text{if } \boldsymbol{x} \neq \boldsymbol{0} \\ \boldsymbol{0} & \text{if } \boldsymbol{x} = \boldsymbol{0}, \end{cases} \tag{1}$$

where $||\boldsymbol{x}|| = \sqrt{x_1^2 + \cdots + x_d^2}$, $\boldsymbol{x} = (x_1, \cdots, x_d)^T \in \mathbb{R}^d$. Note that this multivariate notion of sign vector is nothing but the unit direction vector of $\boldsymbol{x}$ and was used in the literature to construct various statistics based on signs (for detail see, Oja, 1999). A multivariate centred rank function can be defined based on the multivariate sign function as

$$R_{F_n}(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^{n} Sign(\boldsymbol{x} - \boldsymbol{X}_i)$$

where $\boldsymbol{x} \in \mathbb{R}^d$ and $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n \in \mathbb{R}^d$ is a random sample with a common distribution function $F$. We assume throughout this paper that $F$ is absolutely continuous with respect to the Lebesgue measure in $\mathbb{R}^d$. It is easy to note that

1. $||R_{F_n}(\boldsymbol{x})|| < 1$ for all $\boldsymbol{x} \in \mathbb{R}^d$.

2. $R_{F_n}(\boldsymbol{x}) = 0$ means that $\boldsymbol{x}$ is the spatial median of the data $\boldsymbol{X}_1, \boldsymbol{X}_2, \cdots, \boldsymbol{X}_n \in \mathbb{R}^d$.

3. Smaller values of $||R_{F_n}(\boldsymbol{x})||$ imply that $\boldsymbol{x}$ is located more centrally with respect to the data points and larger values of $||R_{F_n}(\boldsymbol{x})||$ imply that $\boldsymbol{x}$ is an extreme point with respect to the data cloud. The direction of the vector $R_{F_n}(\boldsymbol{x})$ suggests the direction in which $\boldsymbol{x}$ is extreme compared to the data cloud.

4. $R_F(\boldsymbol{x}) = E(R_{F_n}(\boldsymbol{x}))$ is an injective function of the multivariate distribution function $F$, see Koltchinskii (1997). Hence $R_F(\boldsymbol{x})$ characterises a multivariate distribution.

We can also note that $R_F(\boldsymbol{x})$ is the inverse function of the multivariate geometric quantile function (Chaudhuri, 1996), $Q(\boldsymbol{u})$, in the sense that $R_F(\boldsymbol{x}) = \boldsymbol{u}$ implies that $Q(\boldsymbol{u}) = \boldsymbol{x}$ and vice-versa. If we define a measure of outlyingness by $||R_F(\boldsymbol{x})||$, then it is easy to verify that this measure of outlyingness is invariant under orthogonal and homogeneous scale transformations (see Serfling, 2004, for some related discussion). Multivariate central rank regions can be defined as

$$C_F(p) = \left\{\boldsymbol{x} : ||R_F(\boldsymbol{x})|| \leq r_F(p)\right\}, \quad 0 \leq p < 1,$$

where $r_F(p)$ is the $p$-th quantile of the distribution of $||R_F(\boldsymbol{X})||$ with $\boldsymbol{X} \in \mathbb{R}^d$ having the distribution function $F$. Note that $P(C_F(p)) = p$. In other words, $C_F(p)$ is the central rank region containing the probability mass $p$. Analogous to the scale curve introduced by Liu et al. (1999), we define a measure of scale as

$$V_F(p) = \text{volume of } C_F(p), \quad 0 \leq p < 1.$$

In the univariate case, $C_F(p)$ is an interval around the median which has a probability mass $p$ and $V_F(p)$ is the length of that interval, which coincides with the measure of spread defined by Balanda and MacGillivray (1990). Shorter intervals suggest smaller spread of the distribution and as $p$ varies from 0 to 1, one can study how the spread is changing with the tails. Similarly, $V_F(p)$ in the multivariate case provides a good measure of scale. For smaller values of $p$, it measures the spread of the data around the spatial median and as $p$ increases, it measures the overall spread of the distribution. A plot of $V_F(p)$ against $p$, named as scale curve, is quite useful in detecting the changes in scale with increasing $p$.

A related scale curve based on spatial ranks has also been considered in Serfling (2002), but the central regions considered there were not indexed by the probability mass $p$. There is another alternative definition of dispersion function based on spatial median developed by Averous and Meste (1997), who extended the univariate interquantile intervals to multivariate "median balls" indexed by their radii, as a class of central regions. Under some regularity conditions on the distribution $F$, the probability mass of a median ball is a nondecreasing function of its radius and that yields a "median ball" analogue of the scale curve defined above, which has not been investigated in detail in the literature.

In Figure 1, we plot the scale curves for standard bivariate normal, Laplace and $t$ distribution with 3 degrees of freedom. We note that the spread of these distributions are similar for the smaller values of $p$, but the scale increases faster for the bivariate Laplace and $t$ distribution compared to the bivariate normal distribution, which suggests that the Laplace and $t$ distributions have larger tails compared to normal. For spherically symmetric distributions with centre of symmetry 0, we know that the length of the random vector $||\boldsymbol{X}||$ and its direction $\boldsymbol{X} = \boldsymbol{X}/||\boldsymbol{X}||$ are independent and the length of the rank vector, $R_F(\boldsymbol{x})$ only depends on the length of the vector $\boldsymbol{x} - \boldsymbol{\theta}$, where $\boldsymbol{\theta}$ is the centre of the spherical symmetry(see Oja, 2010). As a consequence of this result, the central rank region $C_F(p)$ for spherically symmetric random vectors can be written as

$$C_F(p) = \{\boldsymbol{x} : ||\boldsymbol{x} - \boldsymbol{\theta}|| \leq \xi_F(p)\}$$

where $\xi_F(p)$ is the $p$-th quantile of the distribution of $||\boldsymbol{X} - \boldsymbol{\theta}||$. Thus the central rank regions coincide with the level sets of the density of $F$, if the density exists. We can also write down

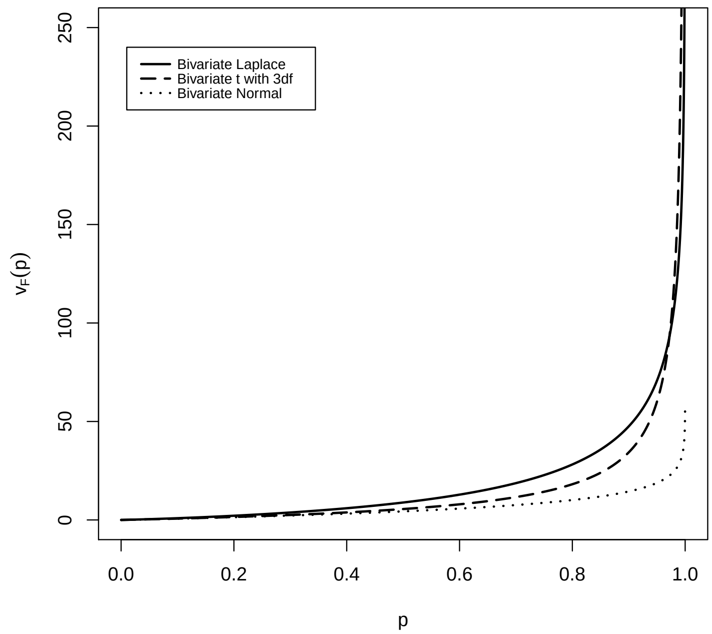


Figure 1: Multivariate scale curves for bivariate normal, bivariate Laplace and bivariate $t$ distribution with 3 degrees of freedom with location $\boldsymbol{\theta} = (0,0)^T$ and scale matrix $\Sigma = \boldsymbol{I}_2$.

a nice closed form formula for the scale curve:

$$V_F(p) = \frac{\pi^{d/2}(\xi_F(p))^d}{\Gamma(\frac{d}{2}+1)}.$$

In a recent study Girard and Stufler (2015a) discussed about the behaviour of extreme geometric quantiles. They showed that if the underlying distribution a random variable has a finite scale matrix, then any extreme geometric quantiles can be estimated accurately. They also discussed that the norm of an extreme geometric quantile is the largest in the direction where the variance is the smallest and outlier detection would be dangerous without a preliminary transformation retransformation procedure. Girard and Stufler (2015b) also discussed extreme geometric quantiles when the integrability conditions are not satisfied. As a major deficiency of the above definition of multivariate rank vector is that they are not

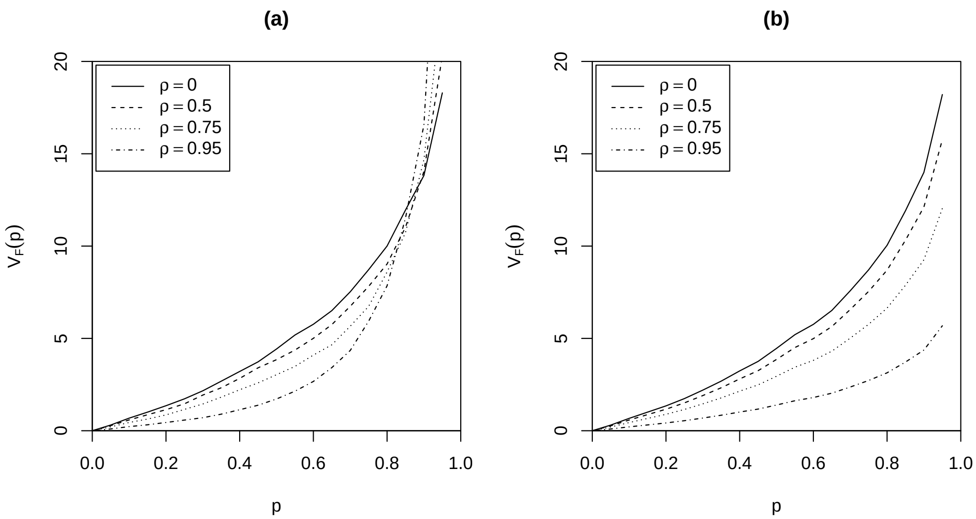


Figure 2: Multivariate scale curves for bivariate normal with mean $\boldsymbol{\theta} = (0,0)^T$ for different values of the correlation coefficient $\rho$.(a) Non-affine equivariant version, (b) affine equivariant version.

invariant under general affine transformations of the data. Consequently, the scale curves based on central rank regions are not affine equivariant. To illustrate, we plot the scale curves of the bivariate normal distributions with mean vector $(0,0)^T$ for different correlation coefficients $\rho$ in Figure 2(a). From the shape of the distributions, we know that the scale decreases with the increasing value of $\rho$, but that behaviour is not reflected in the scale curves. As a remedial measure, we need to make the definition of multivariate ranks to be affine invariant in the sense that if the distribution of $\boldsymbol{X}$ is denoted by $F$ and the distribution of $\boldsymbol{Y} = \mathbf{A}\boldsymbol{X} + \boldsymbol{b}$ is denoted by $G$ for some nonsingular matrix $\mathbf{A}$ and $d$-dimensional vector $\boldsymbol{b}$, then $R_G(\mathbf{A}\boldsymbol{x} + \boldsymbol{b}) = R_F(\boldsymbol{x})$. There are many ways one can achieve that (see Serfling, 2010). In this paper, we adopt the transformation retransformation technique as proposed by Chakraborty (2001).

Let $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ be data points on $\mathbb{R}^d$. Let $\alpha = \{i_0, i_1, \cdots, i_d\}$ denote a set of $(d+1)$ indices. Then we can define a data-driven coordinate system with $\boldsymbol{X}_{i_0}$ as the origin and the coordinate axes given by the vectors $\boldsymbol{X}_{i_1} - \boldsymbol{X}_{i_0}, \ldots, \boldsymbol{X}_{i_d} - \boldsymbol{X}_{i_0}$. Transforming all the observations into the new coordinate system and compute the multivariate rank vector in the new coordinate system by

$$R_{F_n}(\boldsymbol{x}) = \frac{1}{n} \sum_{i=1, i \notin \alpha}^{n} \frac{\{\boldsymbol{X}(\alpha)\}^{-1}(\boldsymbol{x} - \boldsymbol{X}_i)}{|| \{\boldsymbol{X}(\alpha)\}^{-1}(\boldsymbol{x} - \boldsymbol{X}_i)||} ,$$

where $\boldsymbol{X}(\alpha)$ is the $d \times d$ matrix formed with the columns $\boldsymbol{X}_{i_1} - \boldsymbol{X}_{i_0}, \cdots \boldsymbol{X}_{i_d} - \boldsymbol{X}_{i_0}$. The optimal transformation matrix $\boldsymbol{X}(\alpha)$ is obtained by minimising the criterion function

$$t(\alpha) = \frac{\text{trace}[\boldsymbol{X}(\alpha)^T \Sigma^{-1} \boldsymbol{X}(\alpha)]/d}{\{\det[\boldsymbol{X}(\alpha)^T \Sigma^{-1} \boldsymbol{X}(\alpha)]\}^{1/d}} ,$$

where $\Sigma$ is the scale matrix associated with the underlying distribution. For optimality results of such a transformation and related discussion, see Chakraborty (2001). It is easy to observe that the above definition of $R_{F_n}(\boldsymbol{x})$ and the corresponding $R_F(\boldsymbol{x})$ are invariant under general affine transformations. From now on, we restrict our discussion to the affine equivariant versions of the multivariate ranks and corresponding scale curves. The following Theorem states a closed form formula for the affine equivariant scale curve when the underlying distribution is elliptically symmetric.

**Theorem 2.1.** *If the distribution of the random vector $\boldsymbol{X}$ is elliptically symmetric, that is, it has a density of the form*

$$f(\boldsymbol{x}) = |\Sigma|^{-1/2} h((\boldsymbol{x} - \boldsymbol{\theta})^T \Sigma^{-1} (\boldsymbol{x} - \boldsymbol{\theta})) \tag{2}$$

*for some strictly decreasing, continuous, non-negative scalar function $h$ and positive definite matrix $\Sigma$ and $R_F(\boldsymbol{x})$ is the affine invariant spatial rank function as defined before, we have*

$$V_F(p) = \frac{\pi^{\frac{d}{2}} |\Sigma|^{1/2} \zeta_p^d}{\Gamma(\frac{d}{2} + 1)} , \tag{3}$$

*where $P((\boldsymbol{X} - \boldsymbol{\theta})^T \Sigma^{-1} (\boldsymbol{X} - \boldsymbol{\theta}) \leq \zeta_p^2) = p$.*

Note that $\zeta_p$ is invariant under affine transformations and thus it can be computed from the distribution of $||\boldsymbol{Y}||$, where $\boldsymbol{Y}$ has the corresponding spherically symmetric distribution, with location of symmetry 0.

In Figure 2(b), we plot the affine equivariant version of the scale curves based on the affine invariant rank vectors for the bivariate normal distribution with different correlation coefficients $\rho$ and we observe that the ordering between the scale curves with higher values of the correlation coefficient showing smaller scales. So we can conclude that our definition of multivariate rank function is giving an actual picture that the scale decreases with the increasing value of $\rho$ which was not evident in Figure 2(a).

# 3 Scale-Scale Plot

We noted earlier that verifying the distributional assumptions of the multivariate data and comparing samples of multivariate data still remain as difficult tasks. In this section, we propose a scale-scale plot to compare multivariate distributions as a generalization of the univariate quantile quantile plot. If $F$ and $G$ are two $d$-dimensional distributions, we define a scale-scale plot as a plot of $V_G(p)$ against $V_F(p)$, $0 \leq p < 1$, where the functionals $V_F$ and $V_G$ are the volumes of the affine equivariant central rank regions as defined in the previous section. If $F = G$, then the scale-scale plot will be the $45°$ line passing through the origin. Now as the volume $V_F(p)$, depends on the determinant of the scale matrix (see (3)), we may also get a straight line for the scale-scale plot, making an angle of $45°$, when $F \neq G$. Thus in this situation we can detect the shift upto an orthogonal transformation. Since $V_F(0) = 0$ for all continuous distributions $F$, the scale-scale plot will always pass through the origin and we cannot detect a change in origin with the scale-scale plot. However, if we can detect that the two multivariate distributions are same upto a location shift, quite often it is not difficult to estimate the location shift efficiently. For elliptically symmetric distributions $F$ and $G$, we have the following characterization:

**Theorem 3.1.** *Assume that $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{R}^d$ have distributions $F$ and $G$, respectively, which are elliptically symmetric. Then $\boldsymbol{Y} = \mathbf{A}\boldsymbol{X} + \boldsymbol{b}$ for some $d \times d$ matrix $\mathbf{A}$ and $d$-dimensional vector $\boldsymbol{b}$ if and only if $V_G(p) = k.V_F(p)$, $0 \leq p < 1$ for some $k > 0$.*

The above theorem suggests that if $\boldsymbol{X}$ and $\boldsymbol{Y}$ are in the same elliptically symmetric family of distribution but possibly differ in the location parameter $\boldsymbol{\theta}$ and the scale matrix $\Sigma$, the scale-scale plot will be a straight line. The slope of that straight line is determined by the determinants of the scale matrices associated with them. In Figure 3, we present the scale-scale plots comparing the bivariate normal distributions with different correlation coefficients with the standard bivariate normal distribution and observe that the slopes of the straight lines in these scale-scale plots are decreasing with the increasing value of the correlation coefficient $\rho$.

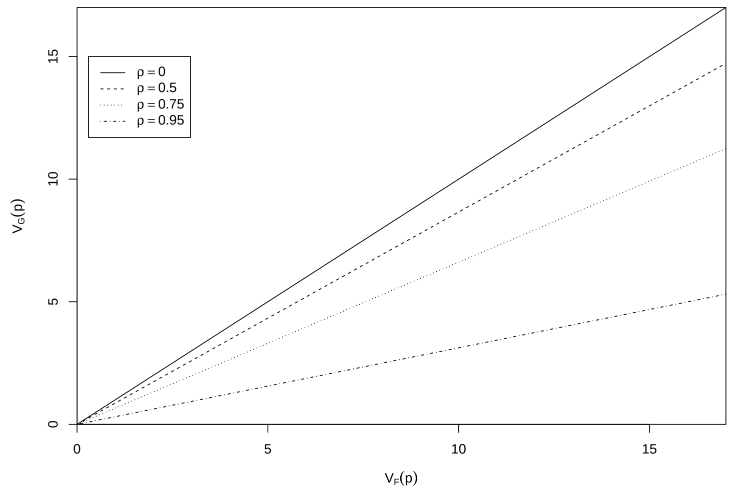


Figure 3: Scale-Scale plots comparing bivariate normal distributions with mean $\boldsymbol{\theta} = (0,0)^T$ and different correlation coefficients $\rho$ with the standard bivariate normal distribution.

All scale-scale plots are strictly increasing, with slope at the point $(V_F(p), V_G(p))$ equal to $V'_G(p)/V'_F(p)$. If $F$ and $G$ are elliptically symmetric, the scale-scale plots can be interpreted in a same way as the quantile-quantile plots for the univariate data, but they can be very different in non-elliptical cases. If all the points $(V_F(p), V_G(p))$ on the scale-scale plot with $p$ near 1 lie above the line joining $(V_F(0.5), V_G(0.5))$ and the origin, $G$ exhibits a greater movement of probability mass from its "shoulders" into its tails than does $F$, that is the tail of $G$ is more "stretched" than $F$. Similarly, if the slope of the tangent to the scale-scale plot at $(V_F(p), V_G(p))$ is greater for $p$ near 0 than the slope of its tangent at the origin, then $G$ exhibits a greater movement of probability mass from its "shoulders" into its centre than does $F$ or in other words the probability mass falls away quickly from its spatial median for $G$ than $F$.

## 3.1 One-Sample Case

The proposed scale-scale plot can be used to check the distributional assumptions of the multivariate data visually as an extension of the univariate quantile quantile plot. We now discuss some application of the proposed Scale-Scale plot for elliptically symmetric distributions. If $F_0$ is the hypothesised distribution function (which is elliptically symmetric) upto a location parameter and $F_n$ denotes the empirical distribution function of the multivariate data $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$, we can make a scale-scale plot by plotting $V_{F_n}(p)$ against $V_{F_0}(p)$. If the plot is close to a straight line, we can conclude that the underlying distribution does not deviate significantly from $F_0$ upto location and scale parameters. For illustration, we use the iris data and check for multivariate normality of the underlying distribution of the three separate species namely iris setosa, iris versicolor and iris virginica. In Figure 4, we take $F_0$ as the bivariate standard normal distribution and observe that all of these plots are close to straight lines except for a few points on the upper tail and justify the assumption of multivariate normality for analysing them. We would like to mention that as these data sets contain only 50 observations each, estimates of the boundary of the extreme rank regions are not very precise and estimates of their volumes also have larger variability and thus deviations from the straight line pattern for few points in the upper end is not unexpected.

In Figure 5, we compare samples from simulated bivariate gamma distributions with the standard bivariate normal distribution. A sample of size $n = 1000$ is simulated from the bivariate gamma density

$$f(x_1, x_2) = \frac{1}{\{\lambda^\alpha \Gamma(\alpha)\}^2} e^{-(x_1+x_2)/\lambda} (x_1 x_2)^{\alpha-1}, x_1, x_2 \geq 0$$

for different values of $\alpha$ and $\lambda$, so that the mean vectors of the distributions remain same. For small value of $\alpha$, bivariate gamma is a highly skewed distribution and their scale-scale plot compared to bivariate normal is a nonlinear curve with scales, $V_G(p)$, increasing sharply with higher values of $p$. However, as $\alpha$ increases, the scale-scale plots become more linear and we observe that the scale-scale plot for the sample from bivariate gamma with $\alpha = 10$ and $\lambda = 0.1$ is very close to the $45°$ line, which is in line with the distributional convergence of the gamma distribution to the normal distribution as $\alpha \to \infty$.

## 3.2 Two Sample Case

We can also use the scale-scale plot to compare two multivariate samples. Suppose $F_n$ and $G_n$ are the empirical distribution functions of the two independent samples, respectively, under the assumption that they are from some elliptically symmetric distribution family. We plot $V_{G_n}(p)$ against $V_{F_n}(p)$ for $0 \leq p < 1$ to construct the scale-scale plot of the two samples.

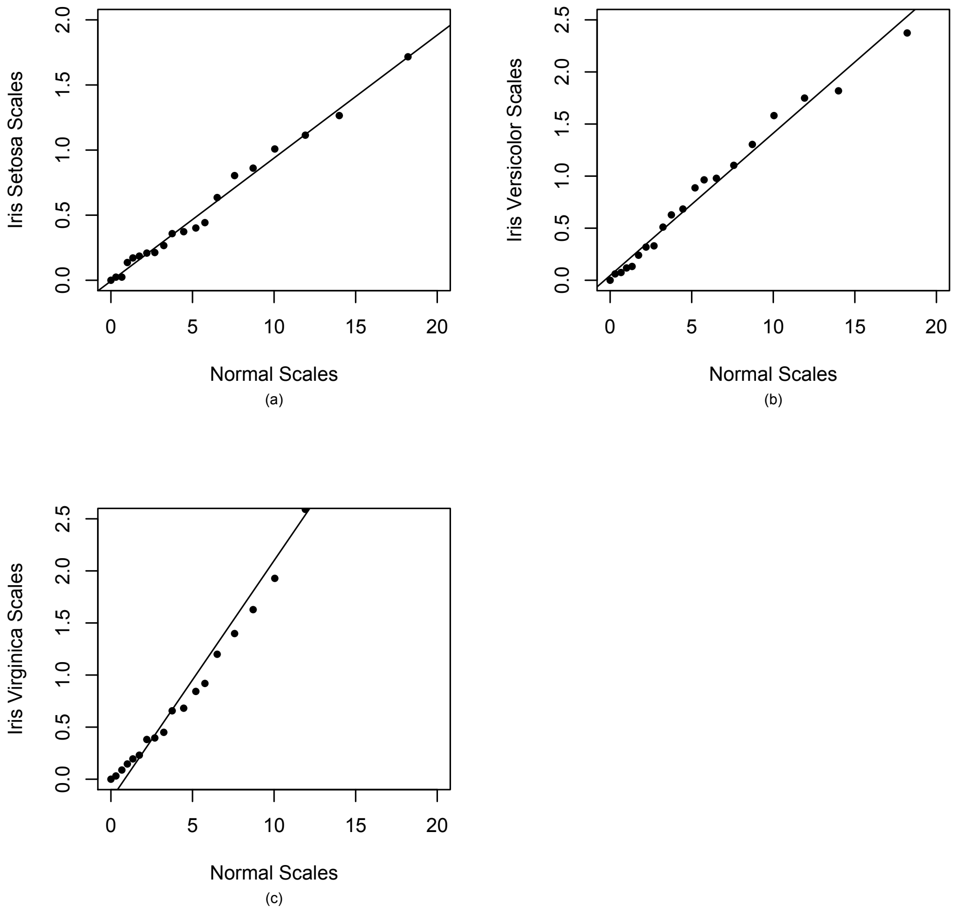


Figure 4: Scale-Scale plot for Iris data to compare with multivariate normal distribution. (a) Iris Setosa, (b) Iris Versicolor, (c) Iris Virginica

The measurements were taken on diameter of rings for the first-year freshwater growth and that for the first-year marine growth for Alaskan and Canadian salmons (See Table 11.2, Johnson and Wichern (2002)). Sample sizes are 50 for both Alaskan-born and Canadian born salmons. This is a nice example on classification techniques used in the literature under the assumption of multivariate normality with the same covariance matrix. To justify, we may use our proposed scale-scale plot in Figure 6, which is almost on a 45° line. That suggests

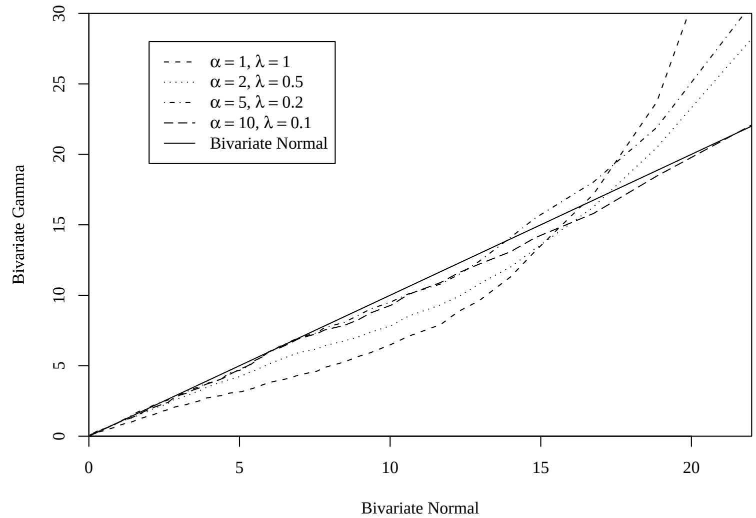


Figure 5: Scale-Scale plots comparing bivariate gamma distribution with the standard bivariate normal distribution.

that the data distribution for both groups of salmons is the same elliptically symmetric distribution with the same scale matrix and possibly with a different location vector.

# 4 Other Applications

In this section, we consider some other applications of the proposed scale-scale plots. It is almost obvious from the construction that the scale-scale plots can be used to detect a change in scale, but in the following we also propose a method to use it to detect a shift in the classical location problem.

## 4.1 Test of Location

In this section, we propose a test of location as an application of the proposed scale-scale plots following the ideas of Singh, Tyler, Zhang, and Mukherjee (2009). Let $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ be a random sample from a $d$-dimensional distribution $F$, which is symmetric around $\boldsymbol{\theta} \in \mathbb{R}^d$ in the sense that $\boldsymbol{X}_i - \boldsymbol{\theta} \stackrel{d}{=} \boldsymbol{\theta} - \boldsymbol{X}_i$. We are interested to test the null hypothesis $H_0$ :

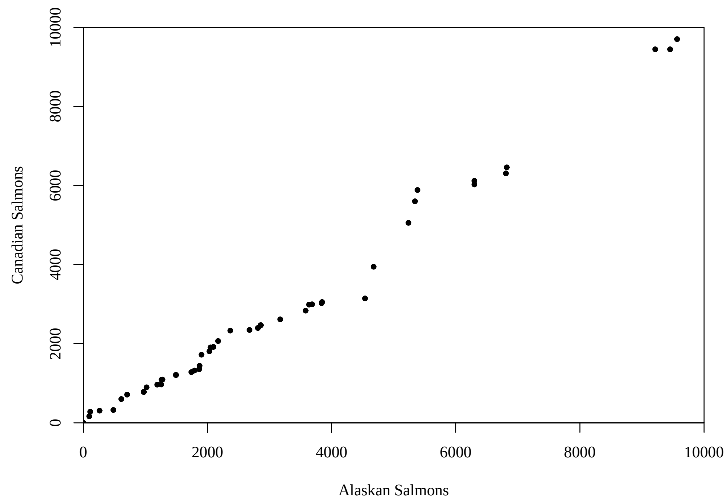


Figure 6: Scale-Scale plot comparing the distributions of Alaskan and Canadian Salmons.

$\boldsymbol{\theta} = \boldsymbol{\theta}_0$ against $H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. Note that under $H_0$, the distributions of $\boldsymbol{X}_i$ and its reflection $2\boldsymbol{\theta}_0 - \boldsymbol{X}_i$ are identical and the scale-scale plot of the combined sample $\{\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n\} \cup \{2\boldsymbol{\theta}_0 - \boldsymbol{X}_1, \cdots, 2\boldsymbol{\theta}_0 - \boldsymbol{X}_i\}$ against the original data $\{\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n\}$ will be nearly a 45° line. However, if the null hypothesis does not hold, the scale of the combined data will be more and the scale-scale plot will move away from the 45° line. Using this principle, we construct a test procedure as follows:

1. Define
$$\boldsymbol{Y}_i = \begin{cases} \boldsymbol{X}_i & \text{with probability } 0.5 \\ 2\boldsymbol{\theta}_0 - \boldsymbol{X}_i & \text{with probability } 0.5, \end{cases}$$
for $i = 1, \cdots, n$.

2. Construct a scale-scale plot of $\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_n$ against $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$.

3. There are $2^n$ possible samples $\{\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_n\}$. However, it is not practical to construct scale-scale plots for all of them for large $n$. One can repeat Steps 1 and 2 for a large number of random subsets $\{\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_n\}$ and construct a band of scale-scale plots.

4. If the 45° line is in the bottom 5% of the band of scale-scale plots or below the band altogether, the null hypothesis is rejected.

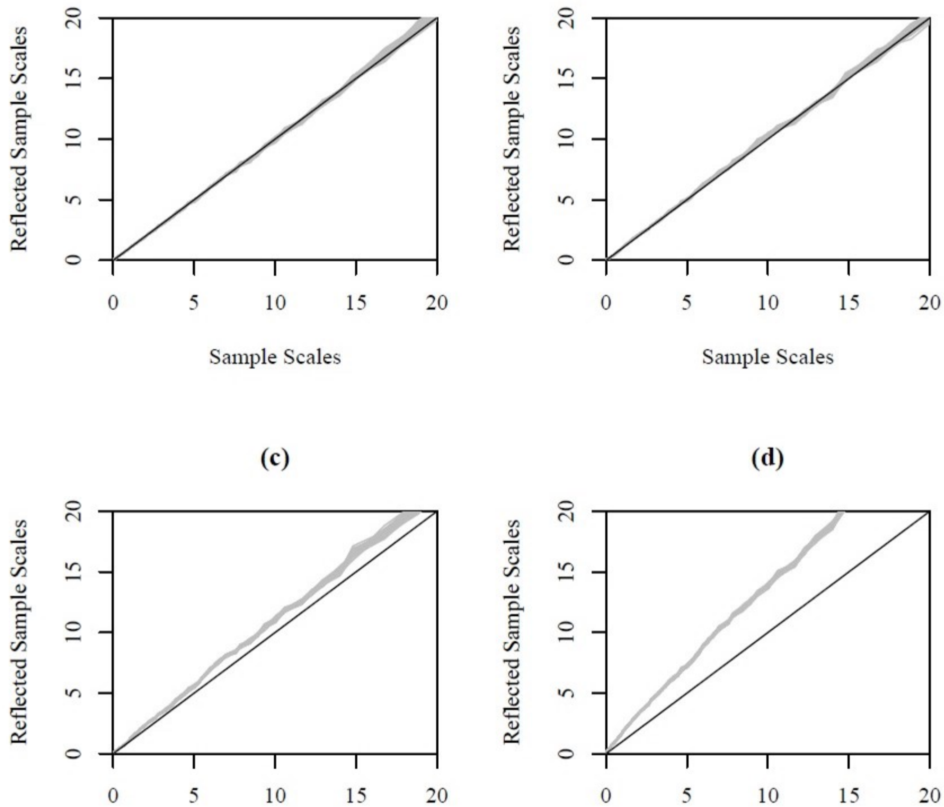


Figure 7: Scale-Scale plot for test of location for bivariate normal distribution for $n = 500$. (a) $\boldsymbol{\theta} = (0, 0)^T$, (b) $\boldsymbol{\theta} = (0.2, 0.2)^T$, (c) $\boldsymbol{\theta} = (0.5, 0.5)^T$, (d) $\boldsymbol{\theta} = (1.0, 1.0)^T$

For illustration, we present plots of the above test procedure in Figure 7 where the data are simulated from a bivariate normal distribution of sample size $n = 500$ with $\Sigma = \mathbf{I}$ and different values of the mean $\boldsymbol{\theta}$. Figure 7(a) is the plot under null hypothesis $H_0 : \boldsymbol{\theta} = \mathbf{0}$ and (b), (c) and (d) present the plots for the alternatives $\boldsymbol{\theta} = (0.2, 0.2)^T$, $(0.5, 0.5)^T$ and $(1.0, 1.0)^T$ respectively. We see that for a small shift in location, the scale-scale plot is not that effective in detecting the shift. However, for moderate to large shifts of the location, this visual tool detects the shift quite effectively.

To formally compute the $p$-values of the proposed test, we have used the proportion of scale-scale plots in the band below the 45° line for some specific values of $p$, for example, at $p = 0.50$, which is the comparison of volumes of the central rank regions containing 50% of the data. Using that criteria, we present a small sample simulation study of the power of the test for different values of $p$.

In this study we compute the power for three elliptically symmetric distributions, bivariate normal, bivariate Laplace and bivariate $t$ with 3 degrees of freedom with the scale matrix $\Sigma = \boldsymbol{I}$ and sample size $n = 30$ and present it in Table 1. The null hypothesis is $H_0 : \boldsymbol{\theta} = (0,0)^T$ and the alternatives are $\boldsymbol{\theta} = (r,r)^T$, for $r = 0.1, 0.2, \cdots, 1.0$. The test is performed based on 100 bands with level of significance 0.05, i.e., if 5 of reflected scale curves at $p$ are below the original sample scale curve at $p$, then we reject the null hypothesis. The size and power of this test are computed based on 1000 simulations for $p = 0.25, 0.50, 0.75$. We observe that the estimated sizes are slightly higher but the powers are nevertheless quite encouraging as they increase with $r$ for every $p$. Hence instead of looking at the entire plot we can perform the test based on one single $p$. Perhaps using $p = 0.5$ might be a good idea as the power increases most rapidly for $p = 0.5$ for all three distributions.

Table 1: Finite sample power of the proposed visual test of location by comparing the scale scale plot at the specified values of $p$ for $n = 30$ and $d = 2$. Here the number of scale-scale plots to construct the band is 100 and the simulation size is 1000 and level of significance $= 5\%$. Powers are computed at $\boldsymbol{\theta} = (r,r)^T$ for different values of $r$.

| | | $r$ | | | |
|---|---|---|---|---|---|
| Distribution | | 0.0 | 0.2 | 0.5 | 1.0 |
| Normal | $p = 0.25$ | 0.057 | 0.114 | 0.369 | 0.878 |
| | $p = 0.50$ | 0.068 | 0.141 | 0.531 | 0.980 |
| | $p = 0.75$ | 0.066 | 0.138 | 0.594 | 0.985 |
| $t$ with 3 d.f. | $p = 0.25$ | 0.071 | 0.113 | 0.384 | 0.867 |
| | $p = 0.50$ | 0.071 | 0.131 | 0.448 | 0.916 |
| | $p = 0.75$ | 0.072 | 0.100 | 0.318 | 0.740 |
| Laplace | $p = 0.25$ | 0.053 | 0.090 | 0.281 | 0.772 |
| | $p = 0.50$ | 0.069 | 0.104 | 0.280 | 0.772 |
| | $p = 0.75$ | 0.057 | 0.071 | 0.229 | 0.618 |

## 4.2 Test of Scale

Suppose $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ and $\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_n$ are random samples from the same family of elliptically symmetric distributions with possibly different location vector $\boldsymbol{\theta}$ and scale matrix $\Sigma$.

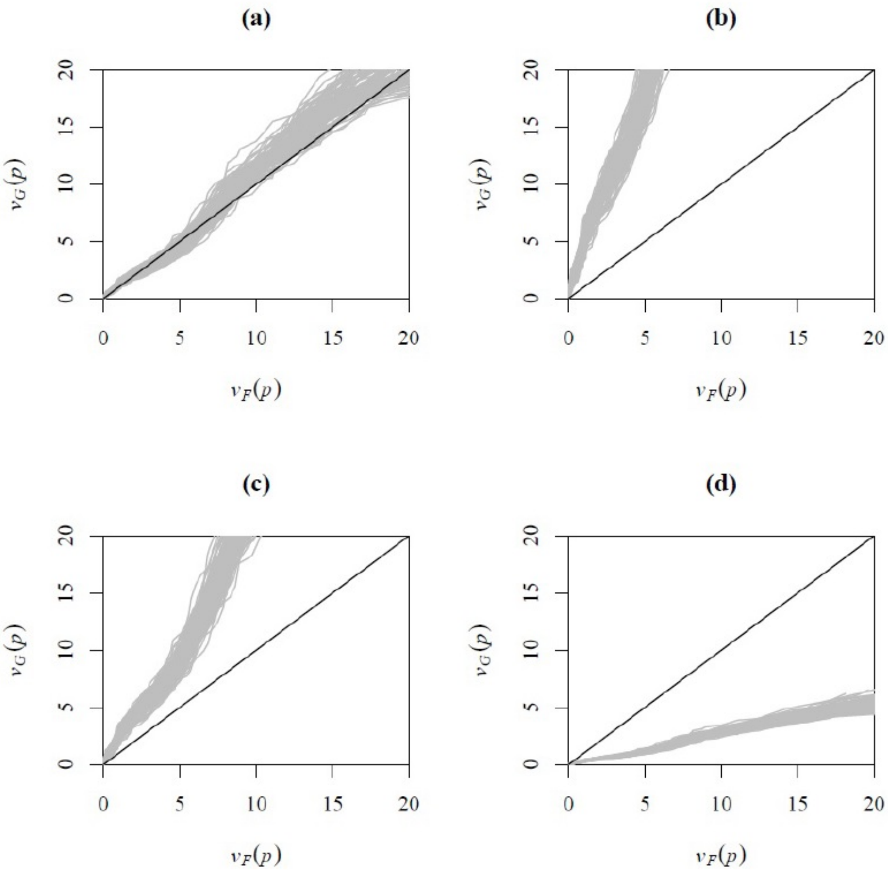


Figure 8: Scale-Scale plot for test of scale for bivariate Laplace distribution for $n = 500$. (a) $\Sigma_Y = \mathbf{I}_2$, (b) $\Sigma_Y = 4\mathbf{I}_2$, (c)$\Sigma_Y = 2\mathbf{I}_2$ and (d) $\Sigma_Y = 0.25\mathbf{I}_2$.

Then, we have $\boldsymbol{Y}_i \stackrel{d}{=} \mathbf{A}\boldsymbol{X}_i + \boldsymbol{b}$ for some $d \times d$ matrix $\mathbf{A}$ and $d \times 1$ vector $\boldsymbol{b}$ and as we noted earlier, the scale-scale plot of such competing samples will be close to a straight line. We would like to test $H_0 : |\Sigma_X| = |\Sigma_Y|$ or alternatively, $H_0 : |\mathbf{A}| = 1$ against $H_1 : |\mathbf{A}| \neq 1$. Under $H_0$, the scale-scale plot will be very close to a 45° line. To construct a graphical test, plot a band of scale-scale plots of $G_n^*$ against $F_n^*$, where $F_n^*$ and $G_n^*$ are bootstrap distributions of $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ and $\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_n$, respectively. Under $H_0$, this band will lie on both sides of the 45° line, whereas under $H_1$, nearly all of it will lie above or below the 45° line. For illustration, we present a few plots in Figure 8. The data are simulated from the bivariate Laplace distribution with mean $\boldsymbol{\theta} = (0, 0)^T$ and $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ have $\Sigma_X = \mathbf{I}_2$. In Figure 8(a), the scale matrix for $\boldsymbol{Y}_1, \cdots, \boldsymbol{Y}_n$ is $\Sigma_Y = \mathbf{I}_2$ and it is $0.25\mathbf{I}_2$, $2\mathbf{I}_2$ and $4\mathbf{I}_2$ in (b), (c) and (d) respectively. We observe that these plots detect the change in scales quite effectively.

## 5 Concluding Remarks

Our proposed scale-scale plot is an effective visual tool to validate distributional assumptions for multivariate data. The scale-scale plot can be constructed based on other depth functions as well but we are using the spatial depth to utilize its nice mathematical properties to illustrate our methods.Under elliptical symmetry of the distributions, we have nice characterisations for the scale-scale plots and they can be used for tests of location in the one sample case or to test equality of the scale in one-sample and two-sample problems. Though we have followed the testing procedure of Singh, Tyler, Zhang, and Mukherjee (2009) for test of location and scale in section 4.1 and 4.2, our plotting method is based on the volumes of the central rank regions; the quantile scale curve proposed by Singh, Tyler, Zhang, and Mukherjee (2009) is the quantile curve of the volumes of random simplices constructed from a multivariate sample. The scale-scale plot is not dependent on the dimension of the data and in principle can be constructed for high dimensional data as well. However, in practice, due to the curse of dimensionality, estimated central rank regions are computationally very unstable for high dimensional data. We have used the `qhull` programme (Barber et al., 1996) to compute the volumes of the central rank regions and that is also not very efficient for very high dimensional data. Though we can construct a scale-scale plot to validate distributional assumptions, it may not be possible to use it for test of location or scale in the case of very high dimensional data with the available algorithms and computing resources.

## A Appendix: Proofs

**Proof of Theorem 2.1:** Let $\boldsymbol{Y} = \Sigma^{-1/2}(\boldsymbol{X} - \boldsymbol{\theta})$, then $\boldsymbol{Y}$ has a spherically symmetric distribution about $\mathbf{0}$. Let $F_0$ denote the distribution of $\boldsymbol{Y}$. Then by the affine invariance of the rank function $R_F(\boldsymbol{x})$, we have $R_{F_0}(\boldsymbol{Y}) = R_{F_0}(\Sigma^{-1/2}(\boldsymbol{X} - \boldsymbol{\theta})) = R_F(\boldsymbol{X})$. This implies that $r_F(p) = r_{F_0}(p)$, where $r_F(p)$ and $r_{F_0}(p)$ are the $p$-th quantiles of the distributions of $||R_F(\boldsymbol{X})||$ and $||R_{F_0}(\boldsymbol{Y})||$ respectively. Thus,

$$\begin{aligned} C_F(p) &= \{\boldsymbol{x} : ||R_F(\boldsymbol{x})|| \le r_F(p)\} \\ &= \left\{\boldsymbol{x} : ||R_{F_0}(\Sigma^{-1/2}(\boldsymbol{x}-\boldsymbol{\theta}))|| \le r_F(p)\right\} \\ &= \left\{\Sigma^{1/2}\boldsymbol{y}+\boldsymbol{\theta} : ||R_{F_0}(\boldsymbol{y})|| \le r_{F_0}(p)\right\} \\ &= \Sigma^{1/2}C_{F_0}(p)+\boldsymbol{\theta}. \end{aligned} \tag{4}$$

Therefore, $V_F(p) = |\Sigma|^{1/2}V_{F_0}(p)$. Now by Theorem 4.3 of Oja(2010), $||R_{F_0}(\boldsymbol{y})||$ is a non-negative increasing function of $||\boldsymbol{y}||$ only, which implies that

$$V_{F_0}(p) = \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2}+1)}\zeta_p^d \tag{5}$$

where $\zeta_p$ is given by,

$$P((\boldsymbol{X}-\boldsymbol{\theta})^T\Sigma^{-1}(\boldsymbol{X}-\boldsymbol{\theta}) \le \zeta_p^2) = P(\mathbf{Y}^T\boldsymbol{Y} \le \zeta_p^2) = p. \tag{6}$$

This proves the theorem. □

**Proof of Theorem 3.1:** Assume that the probability density function of $\boldsymbol{X}$ and $\boldsymbol{Y}$ are given by

$$f(\boldsymbol{x}) = |\Sigma_X|^{-1/2}h_X((\boldsymbol{x}-\boldsymbol{\theta}_X)^T\Sigma_X^{-1}(\boldsymbol{x}-\boldsymbol{\theta}_X))$$

and

$$g(\boldsymbol{y}) = |\Sigma_Y|^{-1/2}h_Y((\boldsymbol{y}-\boldsymbol{\theta}_Y)^T\Sigma_Y^{-1}(\boldsymbol{y}-\boldsymbol{\theta}_Y))$$

respectively. Then by Theorem 2.1,

$$V_F(p) = \frac{\pi^{d/2}|\Sigma_X|^{1/2}\zeta_{X,p}^d}{\Gamma(\frac{d}{2}+1)} \text{ and } V_G(p) = \frac{\pi^{d/2}|\Sigma_Y|^{1/2}\zeta_{Y,p}^d}{\Gamma(\frac{d}{2}+1)},$$

where $\zeta_{X,p}$ and $\zeta_{Y,p}$ are the $p$-th quantiles of the distributions of $\sqrt{(\boldsymbol{X}-\boldsymbol{\theta}_X)^T\Sigma_X^{-1}(\boldsymbol{X}-\boldsymbol{\theta}_X)}$ and $\sqrt{(\boldsymbol{Y}-\boldsymbol{\theta}_Y)^T\Sigma_Y^{-1}(\boldsymbol{Y}-\boldsymbol{\theta}_Y)}$ respectively.

If $\boldsymbol{Y} = \mathbf{A}\boldsymbol{X}+\boldsymbol{b}$, we have $h_X = h_Y$, $\Sigma_Y = \mathbf{A}\Sigma_X\mathbf{A}^T$ and $\boldsymbol{\theta}_Y = \mathbf{A}\boldsymbol{\theta}_X+\boldsymbol{b}$, which implies $(\boldsymbol{X}-\boldsymbol{\theta}_X)^T\Sigma_X^{-1}(\boldsymbol{X}-\boldsymbol{\theta}_X) = (\boldsymbol{Y}-\boldsymbol{\theta}_Y)^T\Sigma_Y^{-1}(\boldsymbol{Y}-\boldsymbol{\theta}_Y)$ and thus $\zeta_{X,p} = \zeta_{Y,p}$ for all $p \in [0,1)$. Therefore, $V_G(p) = |\mathbf{A}|V_F(p)$.

Now to prove the converse, let us assume that $V_G(p) = k.V_F(p)$, then again by Theorem 2.1, we have

$$\zeta_{Y,p} = k^* \zeta_{X,p}, \text{ for some } k^* > 0 \text{ and for all } p \in [0,1).$$

Then by elliptic symmetry of the distributions of $\boldsymbol{X}$ and $\boldsymbol{Y}$, we have

$$\Sigma_Y^{-1/2}(\boldsymbol{Y} - \boldsymbol{\theta}_Y) = k^* \Sigma_X^{-1/2}(\boldsymbol{X} - \boldsymbol{\theta}_X)$$

and therefore

$$\boldsymbol{Y} = k^* \Sigma_Y^{1/2} \Sigma_X^{-1/2} \boldsymbol{X} - k^* \Sigma_Y^{1/2} \Sigma_X^{-1/2} \boldsymbol{\theta}_X + \boldsymbol{\theta}_Y$$

which proves the theorem with

$$\mathbf{A} = k^* \Sigma_Y^{1/2} \Sigma_X^{-1/2} \text{ and } \boldsymbol{b} = -k^* \Sigma_Y^{1/2} \Sigma_X^{-1/2} \boldsymbol{\theta}_X + \boldsymbol{\theta}_Y.$$

□


# Acknowledgements

The authors would like to thank the referee, associate editor and the School of Mathematics, University of Birmingham, UK for support.


BIBLIOGRAPHY


Andrews, D.F., Gnanadesikan, R., and Warner, J.L. (1973), Methods for assessing multivariate normality. In *Multivariate Analysis - III*, (ed. P.R. Krishnaiah), Academic Press, New York, 95–116.

Avérous, J., and Meste, M. (1997), Median balls: an extension of the interquantile intervals to multivariate distributions, *Journal of Multivariate Analysis*, **63**, 222–241.

Balanda, K.P., and MacGillivray, H.L. (1990), Kurtosis and spread, *The Canadian Journal of Statistics*, **18**, 17–30.

Barber, C. B., Dobkin, D. P., Huhdanpaa, H. T. (1996), The Quickhull Algorithm For Convex Hulls, *ACM Transactions on Mathematical Software*, **22**, 469–483.

Barnett, V. (1975), Probability plotting methods and order statistics, *Applied Statistics*, **24**, 95–108.

Bickel, P.J., and Lehmann, E.L. (1979), Descriptive statistics for non-parametric models. IV. Spread. In *Contributions to Statistics, Jaroslav Hájek Memorial Volume* (ed. J. Jureckova), Academia, Prague, 33–40.

Chakraborty, B. (2001), On Affine Equivariant Multivariate Quantiles. *Annals of the Institute of Statistical Mathematics*, **53**, 380–403.

Chaudhuri, P. (1996), On a Geometric Notion of Quantiles for Multivariate Data, *Journal of the American Statistical Association*, **91**, 862–872.

Cleveland, W. (1993), *Visualizing Data*, Hobart Press.

Cook, D.R., Weisberg, S. (1982), *Residuals and Influence in Regression*, Chapman & Hall, New York.

Dhar, S. S., Chakraborty, B., Chaudhuri, P. (2014) Comparison of Multivariate Distributions Using Quantile-Quantile Plots and Related Tests, *Bernoulli*, **20**, 1484-1506.

Easton, G.S., and McCulloch, R.E. (1990), A multivariate generalization of quantile-quantile plots, *Journal of the American Statistical Association*, **85**, 376–386.

Fang, K. T., Kotz, S., Ng. K. W. (1990), *Symmetric Multivariate and Related Distributions.* Chapman & Hall, London.

Girard, S., Stupfler, G. (2015), Intriguing properties of extreme geometric quantiles. REVSTAT-Statistical Journal, to-appear.

Girard, S., Stupfler, G. (2015). Extreme geometric quantiles in a multivariate regular variation framework. *Extremes*, 18(4), 629-663.

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

Healy, M.J.R. (1968), Multivariate normal plotting. *Applied Statistics*, **17**, 157–161.

Johnson, R.A., Wichern, D.W. (2002), *Applied Multivariate Statistical Analysis*, Prentice Hall, New Jersey.

Koltchinskii, V. I. (1997), M-Estimation, Convexity and Quantiles, *The Annals of Statistics*, **25**, 435–477.

Liang, J., and Ng, K.W. (2009), A multivariate normal plot to detect nonnormality, *Journal of Computational and Graphical Statistics*, **18**, 52–72.

Liang, J., Pan, W., and Yang, Z.H. (2004), Characterization-based Q-Q plots for testing multinormality, *Statistics & Probability Letters*, **70**, 183–190.

Liu, R. Y., Parelius, J. M., Singh, K. (1999), Multivariate Analysis by Data Depth: Descriptive Statistics, Graphics and Inference, *The Annals of Statistics*, **27**, 783–858.

Marden, J.I. (1998), Bivariate QQ-plots and spider web plots, *Statistica Sinica*, **8**, 813–826.

Marden, J.I. (2004), Positions and QQ Plots, *Statistical Science*, **19**, 606–614.

Oja, H. (1999), Affine Invariant Multivariate Sign and Rank Tests and Corresponding Estimates: A Review, *Scandinavian Journal of Statistics*, **26** 319–343.

Oja, H. (2010), *Multivariate Nonparametric Methods with R: An Approach Based on Spatial Signs and Ranks*, Springer, New York.

Serfling, R.(2002), A depth function and a scale curve based on spatial quantiles, In *Statistical Data Analysis Based On the $L_1$-Norm and Related Methods*, 25–38. Birkhaser(ed. Y. Dodge)

Serfling, R. (2010), Equivariance and invariance properties of multivariate quantile and related functions, and the role of standardisation, *Journal of Nonparametric Statistics*, Vol. **22**, 915–936.

Singh, K., Tyler, D.E., Zhang, J., and Mukherjee, S. (2009), Quantile scale curves, *Journal of Computational and Graphical Statistics*, **18**, 92–105.

Wang, J., Zhou, W. (2012), A generalized multivariate kurtosis ordering and its applications, *Journal of Multivariate Analysis*, **107**, 169–180.

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

Zuo, Y., and Serfling, R. (2000), Structural properties and convergence results for contours of sample statistical depth functions, *The Annals of Statistics*, **28**, 483–499.