



A Comment on Affine Invariance and Ancillarity in Testing Multivariate Normality

Zbigniew Szkutnik¹

Accepted: 2 February 2021 / Published online: 11 March 2021
© The Author(s) 2021

Abstract

With the aim to clarify some confusion apparently still present in part of the literature on testing for multivariate normality, we collect and discuss some facts on invariance and ancillarity of various variants of, so-called, configuration of a multivariate normal sample. In particular, we show how the choice of the rotation matrix, used in the definition of the configuration, influences its invariance, distribution and ancillarity.

Keywords Multivariate normal distribution · Invariant test · Maximal invariant · Ancillary statistics

Mathematics Subject Classification 62H10 · 62G10 · 62H15

In a recent article by Vexler [10], a test of bivariate normality is proposed. The statistic of that test is a function of the symmetric version of, so-called, sample configuration. Let \mathbf{X} be a (p, n) data matrix with $n > p$. Under the null hypothesis, the columns $\mathbf{X}_1, \dots, \mathbf{X}_n$ of \mathbf{X} form a n -element i.i.d. sample from a p -variate normal distribution $N_p(\mathbf{m}, \mathbf{\Sigma})$, with $\mathbf{m} \in \mathbb{R}^p$ and a positive definite covariance matrix $\mathbf{\Sigma}$. Both \mathbf{m} and $\mathbf{\Sigma}$ are considered unknown. Let $\bar{\mathbf{X}} = n^{-1} \sum_{i=1}^n \mathbf{X}_i$ be the sample mean and $\mathbf{S} = n^{-1}(\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_n^T)(\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_n^T)^T$ be the sample covariance matrix, where $\mathbf{1}_n^T = (1, \dots, 1) \in \mathbb{R}^n$. The symmetric sample configuration, also called the matrix of scaled residuals, is defined as $\mathbf{Z} := \mathbf{S}^{-1/2}(\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_n^T)$, where $\mathbf{S}^{-1/2}$ stands for the symmetric, positive definite square root of the inverse of \mathbf{S} (with $n > p$, \mathbf{S} is almost surely non-singular, cf. [2]). On page 5 of [10], the author describes a Monte Carlo experiment aimed to “experimentally confirm” that the null distribution of the test statistic does not depend on $(\mathbf{m}, \mathbf{\Sigma})$, i.e., that \mathbf{Z} is ancillary. In what follows, we show that, although \mathbf{Z} is not invariant w.r.t. standard groups of data transformations

✉ Zbigniew Szkutnik
szkutnik@agh.edu.pl

¹ Faculty of Applied Mathematics, AGH University of Science and Technology, al. A. Mickiewicza 30, 30-059 Kraków, Poland

usually considered in the context of testing multivariate normality, the null distribution of \mathbf{Z} does not, indeed, depend on $(\mathbf{m}, \boldsymbol{\Sigma})$, so that the Monte Carlo study in [10] was not necessary.

More generally, a sample configuration may also be defined as $\mathbf{C} := \mathbf{L}^{-1}(\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_n^T)$, with any matrix \mathbf{L} such that $\mathbf{S} = \mathbf{L}\mathbf{L}^T$. It is well known that such \mathbf{C} is defined up to left multiplication by a rotation matrix. Although some authors seem to suggest that \mathbf{C} is always ancillary (e.g., [4, Sect. 2]), it is not always true and the distribution and properties of \mathbf{C} depend on the choice of \mathbf{L} .

In what follows, two groups of transformations will be of interest: the group \mathcal{G} of affine transformations $\mathbf{X} \rightarrow \mathbf{A}\mathbf{X} + \mathbf{b}\mathbf{1}_n^T$, with nonsingular (p, p) matrices \mathbf{A} and with $\mathbf{b} \in \mathbb{R}^p$, and its subgroup \mathcal{G}^* , with $\mathbf{A} \in UT(p)$ - the group of upper triangular matrices with positive diagonal. Various questions related to invariant tests for multivariate normality were discussed in [7]. It was shown, in particular, that if $\mathbf{S} = \mathbf{L}\mathbf{L}^T$ with $\mathbf{L} \in UT(p)$, then $\mathbf{B} := \mathbf{L}^{-1}(\mathbf{X} - \bar{\mathbf{X}}\mathbf{1}_n^T)$ is a maximal invariant w.r.t. \mathcal{G}^* and, hence, \mathbf{B} is an ancillary statistic for $(\mathbf{m}, \boldsymbol{\Sigma})$. Additionally, the distribution of \mathbf{B} is invariant w.r.t. left multiplication of \mathbf{B} by fixed, orthogonal (p, p) matrices. This follows directly from [1]. It is shown there (in Sect. 5) that $\mathbf{B} = \mathbf{B}_u\mathbf{D}$, where $\mathbf{B}_u = \mathbf{B}_u(\mathbf{X})$ is a $(p, n-1)$ random matrix, \mathbf{D} is a specific, fixed $(n-1, n)$ matrix and the distribution of \mathbf{B}_u is the Haar measure on the group $SO(n-1)$ of $(n-1, n-1)$ orthogonal matrices, marginalized to the first p rows. This distribution is clearly invariant w.r.t. left multiplication of \mathbf{B}_u by orthogonal (p, p) matrices, say \mathbf{R} , because this corresponds to left multiplication in $SO(n-1)$ by orthogonal, block diagonal matrices $\text{diag}(\mathbf{R}, \mathbf{I}_{n-1-p})$, where \mathbf{I}_{n-1-p} stands for the identity matrix.

Any configuration \mathbf{C} is related to \mathbf{B} through $\mathbf{C} = \mathbf{M}\mathbf{B}$, with an orthogonal matrix \mathbf{M} . Invariance and ancillarity of \mathbf{C} depend on the way \mathbf{M} is defined as a function of \mathbf{X} .

If \mathbf{M} is stochastically independent of \mathbf{B} , then the invariance of the distribution of \mathbf{B} w.r.t. left multiplication by fixed, orthogonal matrices leads via a standard conditioning argument to the conclusion, that the distributions of \mathbf{C} and \mathbf{B} are identical. As an example, with $\mathbf{L} \in UT(p)$, since $\mathbf{S} = \mathbf{L}\mathbf{L}^T = \mathbf{S}^{1/2}\mathbf{S}^{1/2}$, the symmetric configuration satisfies $\mathbf{Z} = \mathbf{M}\mathbf{B}$, with an orthogonal matrix $\mathbf{M} = \mathbf{S}^{-1/2}\mathbf{L}$. Since \mathbf{M} is a function of \mathbf{S} only (because both \mathbf{L} and $\mathbf{S}^{1/2}$ are) and $(\bar{\mathbf{X}}, \mathbf{S})$ is sufficient and complete in the Gaussian model, \mathbf{M} and \mathbf{B} are stochastically independent by the Basu theorem, and \mathbf{Z} has the same distribution as \mathbf{B} and is, hence, ancillary. The same conclusion holds true, if \mathbf{M} is any function of $(\bar{\mathbf{X}}, \mathbf{S})$.

If \mathbf{M} is a function of \mathbf{B} only, then $\mathbf{C} = \mathbf{M}\mathbf{B}$ is clearly ancillary as a function of the ancillary statistic \mathbf{B} , with a distribution possibly different from that of \mathbf{B} . \mathbf{C} is then invariant w.r.t. \mathcal{G}^* , but not necessarily w.r.t. \mathcal{G} .

If, moreover, \mathbf{M} is a function of $\mathbf{B}\mathbf{B}^T$, i.e., of the Mahalanobis distances and angles, which is a maximal invariant w.r.t. \mathcal{G} (see, e.g., [3]), then $\mathbf{C} = \mathbf{M}\mathbf{B}$ is also invariant w.r.t. \mathcal{G} . For an example, see, e.g., [7, Th. 3]. As one of the columns of the configuration \mathbf{C} constructed there is always proportional to $(1, 0, \dots, 0)^T$, its distribution is clearly different from that of \mathbf{B} .

Finally, if \mathbf{M} depends on \mathbf{X} in an arbitrary way, i.e., not necessarily through \mathbf{B} only, then $\mathbf{C} = \mathbf{M}\mathbf{B}$ does not have to be ancillary. \mathbf{M} can be, e.g., a rotation that makes the first column \mathbf{B}_1 of \mathbf{B} parallel to \mathbf{X}_1 , randomly chosen from the Haar

distribution on $\mathcal{SO}(p)$, conditionally on \mathbf{MB}_1 being parallel to \mathbf{X}_1 . Then, clearly, the expectation of the first column of \mathbf{C} is proportional to \mathbf{m} , and \mathbf{C} is not ancillary.

It should be noted that many tests for multivariate normality are based on the symmetric sample configuration \mathbf{Z} and not all of them are \mathcal{G} invariant, even if a strong case for this property is sometimes made, e.g., at the beginning of Sect. 2 in [5]. Invariance w.r.t. \mathcal{G}^* only may also be of interest, if directed tests for multivariate normality against some restricted alternatives are considered, as in [6, 7], or if the goal is maximin testing between some neighborhoods of transformation families of distributions (see, e.g., [8]).

If the sample covariance matrix \mathbf{S} is replaced in the definition of the sample configuration with another affine equivariant estimator $\mathbf{V}(\mathbf{X})$ that satisfies $\mathbf{V}(\mathbf{AX} + \mathbf{b}\mathbf{1}_n^T) = \mathbf{AV}(\mathbf{X})\mathbf{A}^T$ for nonsingular \mathbf{A} and $\mathbf{b} \in \mathbb{R}^p$, e.g., a robust estimator studied in [9], then \mathbf{B} remains a maximal invariant w.r.t. \mathcal{G}^* , but the rows of \mathbf{B} are not orthogonal, as it was the case with \mathbf{S} . As the distribution of \mathbf{B} does not have to be, in that case, invariant under left multiplication by orthogonal matrices, the discussion of the distributional issues does not carry over to that more general case.

Declarations

Conflict of interest The author declares that he has no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Ćmiel A, Szkutnik Z (1991) On the distribution of a useful maximal invariant. *Prob Math Stat* 12:57–65
2. Eaton MR, Perlman MD (1973) The non-singularity of generalized sample covariance matrices. *Ann Stat* 1:710–717
3. Fattorini L (2001) On the assessment of multivariate normality. In: *Atti della XL Riunione Scientifica della Societa Italiana di Statistica*, Firenze, 26-28 Aprile 2001, pp. 313–324
4. Fattorini L, Pisani C (2000) Assessing multivariate normality on the “worst” sample configuration. *Metron* 58:23–38
5. Henze N (2002) Invariant tests for multivariate normality: a critical review. *Stat Pap* 43:467–506
6. Majerski P, Szkutnik Z (2010) Approximations to most powerful invariant tests for multinormality against some irregular alternatives. *Test* 19:113–130
7. Szkutnik Z (1987) On invariant tests for multidimensional normality. *Prob Math Stat* 8:1–10
8. Szkutnik Z (1992) Special capacities, the Hunt–Stein theorem and transformation groups. *Ann Stat* 20:1120–1128
9. Tyler DE (1987) A distribution-free estimator of multivariate scatter. *Ann Stat* 15:234–251

10. Vexler A (2020) Univariate likelihood projections and characterizations of the multivariate normal distribution. *J Multivar Anal* 179:104643

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.