Article

On a Vector-Valued Measure of Multivariate Skewness

Nicola Loperfido

Dipartimento di Economia, Società e Politica, Università degli Studi di Urbino “Carlo Bo”, Via Saffi 42, 61029 Urbino, Italy; nicola.loperfido@uniurb.it

Abstract: The canonical skewness vector is an analytically simple function of the third-order, standardized moments of a random vector. Statistical applications of this skewness measure include semiparametric modeling, independent component analysis, model-based clustering, and multivariate normality testing. This paper investigates some properties of the canonical skewness vector with respect to representations, transformations, and norm. In particular, the paper shows its connections with tensor contraction, scalar measures of multivariate kurtosis and Mardia’s skewness, the best-known scalar measure of multivariate skewness. A simulation study empirically compares the powers of tests for multivariate normality based on the squared norm of the canonical skewness vector and on Mardia’s skewness. An example with financial data illustrates the statistical applications of the canonical skewness vector.

Keywords: kurtosis; multivariate normality testing; skewness; star product; tensor contraction

1. Introduction

Let $\mathbf{x}$ be a p -dimensional random vector with mean $\boldsymbol{\mu}$, nonsingular covariance matrix $\boldsymbol{\Sigma}$ and finite third-order moments. Ref. [1] introduced the vector-valued measure of multivariate skewness as follows:

$$\gamma_{1V} = E(\mathbf{z}^T \mathbf{z} \mathbf{z}),$$

where $\mathbf{z} = \boldsymbol{\Sigma}^{-1/2}(\mathbf{x} - \boldsymbol{\mu})$ is the standardization of $\mathbf{x}$ and $\boldsymbol{\Sigma}^{-1/2}$ is the positive definite symmetric square root of the concentration matrix $\boldsymbol{\Sigma}^{-1}$, that is, the inverse of $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma}^{-1/2} > 0, \boldsymbol{\Sigma}^{-1/2} = \left(\boldsymbol{\Sigma}^{-1/2}\right)^T, \boldsymbol{\Sigma}^{-1/2} \boldsymbol{\Sigma}^{-1/2} = \boldsymbol{\Sigma}^{-1}, \boldsymbol{\Sigma}^{-1/2} \boldsymbol{\Sigma} \boldsymbol{\Sigma}^{-1/2} = \mathbf{I}_p.$$

The i -th element of γ_{1V} is as follows:

$$E(\mathbf{z}^T \mathbf{z} Z_i) = E(Z_1^2 Z_i) + \dots + E(Z_p^2 Z_i),$$

where Z_k is the k -th component of $\mathbf{z}$. We shall refer to γ_{1V} as the canonical skewness vector to distinguish it from less known vector-valued measures of multivariate skewness [2,3]. In the following, all vectors are regarded as column vectors.

The intuition behind γ_{1V} might be better appreciated by considering some special but relevant cases. In the univariate case, that is, when the only component of $\mathbf{x}$ is the random variable X , the skewness vector coincides with the skewness of X , that is, its third standardized moment, as follows:

$$\gamma_{1V}(\mathbf{x}) = \gamma_1(X) = \frac{E\{(X - \mu)^3\}}{\sigma^3},$$

where μ and σ are the expected value and the standard deviation, respectively, of X . The canonical skewness vector is a null vector when $\mathbf{x}$ is centrally symmetric, that is, if $\mathbf{x} - \boldsymbol{\mu}$

Citation: Loperfido, N. On a Vector-Valued Measure of Multivariate Skewness. *Symmetry* **2021**, *13*, 1817. <https://doi.org/10.3390/sym13101817>

Academic Editor: Carlo Cattani

Received: 26 August 2021

Accepted: 22 September 2021

Published: 29 September 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

and if $\mu - \mathbf{x}$ are identically distributed [4]. In the bivariate case, the skewness vector admits the simpler representation as follows:

$$\gamma_{1V} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} E(Z_1^3) + E(Z_1^2 Z_2) \\ E(Z_2^3) + E(Z_2^2 Z_1) \end{pmatrix}.$$

The canonical skewness vector γ_{1V} appears in several areas of multivariate statistical analysis. In model-based clustering, the projection of the standardized data onto the direction of the sample counterpart of γ_{1V} is used to estimate the best discriminant projection [5]. In independent component analysis (ICA), γ_{1V} is the product of an orthogonal matrix and the vector whose i -th element is the skewness of the i -th independent component ([6]). In the semiparametric model posed by [7], γ_{1V} is instrumental in identifying the parameter of the model. Within an invariant coordinate selection (ICS) approach, the vector γ_{1V} might be regarded as the standardized difference between two appropriately chosen random vectors ([8]).

Unfortunately, the canonical skewness vector might be a null vector, even if the underlying distribution is skewed with finite third moments and a positive definite covariance matrix. For example, let the density function of the random vector $\mathbf{x} = (X_1, X_2, X_3)^\top$ be $f(x_1, x_2, x_3; \theta) = 2\phi(x_1)\phi(x_2)\phi(x_3)\Phi(\theta x_1 x_2 x_3)$, where $\phi(\cdot)$ is the standard normal probability density function, $\Phi(\cdot)$ is the standard normal cumulative distribution function, and θ is a non-null real value. Then, $\mathbf{x}$ is a standard random vector, and its only non-null third moment is $E(X_1 X_2 X_3)$. The canonical skewness vector is then the three-dimensional null vector.

The partial skewness $\beta_{1P}(\mathbf{x})$ of $\mathbf{x}$ is just the squared norm of its canonical skewness vector:

$$\beta_{1P}(\mathbf{x}) = \|\gamma_{1V}(\mathbf{x})\|^2 = \gamma_{1V}^\top(\mathbf{x})\gamma_{1V}(\mathbf{x}) = \sum_{i=1}^p \sum_{h=1}^p \sum_{k=1}^p E(Z_h^2 Z_i) E(Z_k^2 Z_i).$$

It was independently proposed by several authors (see, for example, refs. [1,8,9]) as a scalar measure of multivariate skewness; its name reminds that it does not depend on the cross-product moment $E(Z_i Z_j Z_k)$ when i, j , and k differ from each other [6]. Partial skewness is nonnegative, equals zero if $\mathbf{x}$ is centrally symmetric, and is invariant with respect to affine, nonsingular transformations. Its statistical applications include multivariate normality testing [10] and multivariate analysis of variance [9].

Ref. [11] proposed to measure the skewness of $\mathbf{x}$ by the following expectation:

$$\beta_{1T}(\mathbf{x}) = E\left\{\left(\mathbf{z}^\top \mathbf{w}\right)^3\right\},$$

where $\mathbf{z}$ and $\mathbf{w}$ are identically distributed and mutually independent. In the following, we shall refer to β_{1T} as the total skewness since it depends on all third-order standardized moments of $\mathbf{x}$. Just like partial skewness, total skewness is nonnegative, equals zero if the underlying distribution is centrally symmetric, and is invariant with respect to nonsingular affine transformations. Moreover, just like partial skewness, total skewness is used in multivariate normality testing [11] and in multivariate analysis of variance [9]. However, β_{1T} is by far more popular than β_{1P} , to the point of being a default measure of multivariate skewness, as remarked by [3].

The total skewness is always positive when some third-order cumulants of the underlying distribution are non-null and the covariance matrix is positive definite. This feature constitutes a major advantage of total skewness over partial skewness. For example, the total skewness and the partial skewness of the random vector $\mathbf{x} = (X_1, X_2, X_3)^\top$ with probability density function $f(x_1, x_2, x_3; \theta) = 2\phi(x_1)\phi(x_2)\phi(x_3)\Phi(\theta x_1 x_2 x_3)$ are $\beta_{1T}(\mathbf{x}) = 6E^2(X_1 X_2 X_3)$ and $\beta_{1P}(\mathbf{x}) = 0$. As remarked before, the sample counterparts of both skewness measures were used to test normality, but they are suited for testing symmetry since

there are many multivariate distributions which are centrally symmetric without being normal, such as the multivariate Student t distribution.

There are many more measures of multivariate skewness (see, for example, ref. [12]). Similarly, there are other notions of multivariate symmetry other than central symmetry, such as weak symmetry, sign symmetry and elliptical symmetry. A detailed investigation of notions of multivariate symmetry and measures of multivariate skewness falls outside the scope of the present paper. We defer the interested reader to [4], where the previous literature on both topics is thoroughly reviewed.

This paper contributes to the literature on the canonical skewness vector both with empirical and theoretical results. The theorems in Section 2 provide some alternative representations for γ_{1V} , β_{1P} and β_{1T} , together with some skewness–kurtosis inequalities and some insights into the behavior of these skewness measures under some well-known transformations. The simulation studies in Section 3 compare the performance of partial skewness and total skewness for testing multivariate symmetry. Section 4 uses financial data to illustrate the statistical applications of the canonical skewness vector and partial skewness. Section 5 contains some concluding remarks and hints for future research. The proofs of the theorems are relegated to Appendix A.

2. Theory

This section contains several theoretical results related to the canonical skewness vector. Theorems 1 and 2 represent the canonical skewness vector by means of the star product of matrices and the vectorization operator. Theorems 3 and 4 investigate the behavior of the canonical skewness vector under linear transformations and convolution. Theorems 5 and 6 focus on partial skewness and total skewness by providing alternative representations and by establishing skewness–kurtosis inequalities.

The third moment matrix of a p -dimensional random vector $\mathbf{x}$ with finite third-order moments is as follows:

$$\mathbf{M}_{3,\mathbf{x}} = \mathbb{E}(\mathbf{x} \otimes \mathbf{x}^\top \otimes \mathbf{x}).$$

The matrix $\mathbf{A} \otimes \mathbf{B}$ denotes the Kronecker product of the matrices $\mathbf{A} = \{a_{ij}\} \in \mathbb{R}^p \times \mathbb{R}^q$ and $\mathbf{B} = \{b_{ij}\} \in \mathbb{R}^h \times \mathbb{R}^k$, that is, the block matrix whose i, j -th block is the matrix $a_{ij}\mathbf{B}$. The third cumulant of $\mathbf{x}$, that is, the third moment of $\mathbf{x} - \boldsymbol{\mu}$, is the following:

$$\mathbf{K}_{3,\mathbf{x}} = \mathbf{M}_{3,\mathbf{x} - \boldsymbol{\mu}} = \mathbb{E}\{(\mathbf{x} - \boldsymbol{\mu}) \otimes (\mathbf{x} - \boldsymbol{\mu})^\top \otimes (\mathbf{x} - \boldsymbol{\mu})\}.$$

The third standardized moment of $\mathbf{x}$, which coincides with its third standardized cumulant, is the third moment (cumulant) of $\mathbf{z}$, given as follows:

$$\mathbf{K}_{3,\mathbf{z}} = \mathbf{M}_{3,\mathbf{z}} = \mathbb{E}(\mathbf{z} \otimes \mathbf{z}^\top \otimes \mathbf{z}).$$

The matrices $\mathbf{M}_{3,\mathbf{x}}$, $\mathbf{K}_{3,\mathbf{x}}$ and $\mathbf{K}_{3,\mathbf{z}}$ are the matricized versions of the third moment tensor, as follows:

$$\mathcal{M}_{3,\mathbf{x}} = \{\mathbb{E}(X_i X_j X_h)\},$$

the third cumulant tensor, as follows:

$$\mathcal{K}_{3,\mathbf{x}} = \{\mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)(X_h - \mu_h)]\}$$

and the third standardized moment tensor, as follows:

$$\mathcal{M}_{3,\mathbf{z}} = \{\mathbb{E}(Z_i Z_j Z_h)\},$$

that is, the third-order arrays containing all the third moments of $\mathbf{x}$, $\mathbf{x} - \boldsymbol{\mu}$, and $\mathbf{z}$. Third-order tensors provide a natural tool for representing the skewness of a random vector [1]. In

particular, the canonical skewness vector is just the tensor contraction of $\mathbf{K}_{3,\mathbf{z}}$ and $\mathbf{I}_p$. Its i -th component is as follows:

$$\sum_{j=1}^p \sum_{h=1}^p \mathbb{E}(Z_i Z_j Z_h) \delta_{jh}, \text{ where } \delta_{jh} = \begin{cases} 1 & j = h \\ 0 & j \neq h \end{cases}.$$

Some tensor contractions might be represented by means of the star product of matrices as defined by [13]. The star product of the $a \times b$ matrix $\mathbf{M} = \{m_{ij}\}$ and the $ac \times bd$ block matrix $\mathbf{N} = \{\mathbf{N}_{ij} \in \mathbb{R}^c \times \mathbb{R}^d\}$ is the $c \times d$ matrix as follows:

$$\mathbf{M} \star \mathbf{N} = \sum_{i=1}^a \sum_{j=1}^b m_{ij} \mathbf{N}_{ij},$$

that is, the linear combination of the blocks of $\mathbf{N}$ where the coefficients are the elements of $\mathbf{M}$. The following theorem shows that the canonical skewness vector is the star product of an identity matrix and the third standardized moment.

Theorem 1. Let $\mathbf{z}$ be the standardization of a p -dimensional random vector $\mathbf{x}$ with a positive definite covariance matrix and finite third-order moments. Then, the canonical skewness vector γ_{1V} of $\mathbf{x}$ is the star product of the $p \times p$ identity matrix $\mathbf{I}_p$ and the third standardized cumulant matrix $\mathbf{K}_{3,\mathbf{z}}$:

$$\gamma_{1V} = \mathbf{I}_p \star \mathbf{K}_{3,\mathbf{z}}.$$

We illustrate the above theorem with the bivariate random vector $(X_1, X_2)^\top$, whose standardization is $(Z_1, Z_2)^\top$. The star product of the 2×2 identity matrix and the third standardized moment of $(X_1, X_2)^\top$ is as follows:

$$\begin{aligned} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \star \begin{pmatrix} \mathbb{E}(Z_1^3) & \mathbb{E}(Z_1^2 Z_2) \\ \mathbb{E}(Z_1^2 Z_2) & \mathbb{E}(Z_1 Z_2^2) \\ \mathbb{E}(Z_1^2 Z_2) & \mathbb{E}(Z_1 Z_2^2) \\ \mathbb{E}(Z_1 Z_2^2) & \mathbb{E}(Z_2^3) \end{pmatrix} = 1 \cdot \begin{pmatrix} \mathbb{E}(Z_1^3) \\ \mathbb{E}(Z_1^2 Z_2) \end{pmatrix} + \\ & 0 \cdot \begin{pmatrix} \mathbb{E}(Z_1^2 Z_2) \\ \mathbb{E}(Z_1 Z_2^2) \end{pmatrix} + 0 \cdot \begin{pmatrix} \mathbb{E}(Z_1^2 Z_2) \\ \mathbb{E}(Z_1 Z_2^2) \end{pmatrix} + 1 \cdot \begin{pmatrix} \mathbb{E}(Z_1 Z_2^2) \\ \mathbb{E}(Z_2^3) \end{pmatrix} = \\ & \begin{pmatrix} \mathbb{E}(Z_1^3) \\ \mathbb{E}(Z_1^2 Z_2) \end{pmatrix} + \begin{pmatrix} \mathbb{E}(Z_1 Z_2^2) \\ \mathbb{E}(Z_2^3) \end{pmatrix} = \begin{pmatrix} \mathbb{E}(Z_1^3) + \mathbb{E}(Z_1 Z_2^2) \\ \mathbb{E}(Z_1^2 Z_2) + \mathbb{E}(Z_2^3) \end{pmatrix}, \end{aligned}$$

which coincides with the canonical skewness vector of $(X_1, X_2)^\top$, as defined in the Introduction.

A location vector $\mathbf{l}(\mathbf{x})$ and a scatter matrix $\mathbf{S}(\mathbf{x})$ of a p -dimensional random vector $\mathbf{x}$ are a p -dimensional vector and a symmetric, positive definite $p \times p$ matrix, satisfying the following:

$$\mathbf{l}(\mathbf{A}\mathbf{x} + \mathbf{b}) = \mathbf{A}\mathbf{l}(\mathbf{x}) + \mathbf{b} \text{ and } \mathbf{S}(\mathbf{A}\mathbf{x} + \mathbf{b}) = \mathbf{A}\mathbf{S}(\mathbf{x})\mathbf{A}^\top$$

for any $p \times p$ matrix $\mathbf{A}$ of full rank and for any p -dimensional real vector $\mathbf{b}$ (see, for example, ref. [14]). The mean $\boldsymbol{\mu}$ and

$$\boldsymbol{\mu}_{3,\mathbf{x}} = \mathbb{E}\left\{(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\mathbf{x}\right\}$$

are examples of location vectors [8]. The covariance and

$$\mathbf{Q}_x = \mathbb{E}\left\{(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\mathbf{x}\mathbf{x}^\top\right\}$$

are examples of scatter matrices [8]. Let $\mathbf{l}_1(\mathbf{x})$, $\mathbf{l}_2(\mathbf{x})$ and $\mathbf{S}(\mathbf{x})$ be two location vectors and a scatter matrix of a p -dimensional random vector $\mathbf{x}$. Ref. [14] suggests using the following:

$$\|\mathbf{a}(\mathbf{x})\|^2 = \{\mathbf{l}_1(\mathbf{x}) - \mathbf{l}_2(\mathbf{x})\}^\top \mathbf{S}^{-1}(\mathbf{x}) \{\mathbf{l}_1(\mathbf{x}) - \mathbf{l}_2(\mathbf{x})\}$$

as a scalar measure of multivariate skewness, where $\mathbf{S}^{-1}(\mathbf{x})$ is the inverse of $\mathbf{S}(\mathbf{x})$. It measures the skewness in the direction of the vector-valued measure of multivariate skewness:

$$\mathbf{a}(\mathbf{x}) = \mathbf{S}^{-1/2}(\mathbf{x}) \{\mathbf{l}_1(\mathbf{x}) - \mathbf{l}_2(\mathbf{x})\},$$

where $\mathbf{S}^{-1/2}(\mathbf{x})$ is the positive definite symmetric square root of $\mathbf{S}^{-1}(\mathbf{x})$. The following theorem represents the canonical skewness vector and the partial skewness as $\mathbf{a}(\mathbf{x})$ and $\|\mathbf{a}(\mathbf{x})\|^2$, where the following holds:

$$\mathbf{l}_1(\mathbf{x}) = \boldsymbol{\mu}_{3,x}, \mathbf{l}_2(\mathbf{x}) = p\boldsymbol{\mu} \text{ and } \mathbf{S}(\mathbf{x}) = \boldsymbol{\Sigma}.$$

The theorem also represents the canonical skewness vector by means of simple matrix functions acting on the third cumulant and on the covariance matrix.

Theorem 2. Let $\mathbf{x}$ be a p -dimensional random vector with a positive, definite covariance matrix and finite third-order moments. Additionally, let $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}^{-1}$, and $\mathbf{K}_{3,x}$ be the mean vector, the concentration matrix and the third cumulant of $\mathbf{x}$. Then, the canonical skewness vector of $\mathbf{x}$ is as follows:

$$\gamma_{1V} = \boldsymbol{\Sigma}^{-1/2} \mathbf{K}_{3,x}^\top \text{vec}(\boldsymbol{\Sigma}^{-1}) = \boldsymbol{\Sigma}^{-1/2} \left[\mathbb{E}\left\{(\mathbf{x} - \boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\mathbf{x}\right\} - p\boldsymbol{\mu} \right],$$

where $\boldsymbol{\Sigma}^{-1/2}$ is the positive definite symmetric square root of $\boldsymbol{\Sigma}^{-1}$.

The canonical skewness vector is location invariant: $\gamma_{1V}(\mathbf{x}) = \gamma_{1V}(\mathbf{x} + \mathbf{v})$ for any p -dimensional real vector $\mathbf{v}$. On the other hand, the canonical skewness vector is neither invariant nor equivariant with respect to linear transformations: $\gamma_{1V}(\mathbf{A}\mathbf{x})$ may differ from both $\gamma_{1V}(\mathbf{x})$ and $\mathbf{A}\gamma_{1V}(\mathbf{x})$, where $\mathbf{A}$ is a nonsingular $p \times p$ real matrix. Ref. [5] conjectures that $\gamma_{1V}(\mathbf{A}\mathbf{x}) = \mathbf{U}\gamma_{1V}(\mathbf{x})$ for some orthogonal $p \times p$ matrix $\mathbf{U}$. The following theorem makes this statement more precise.

Theorem 3. Let $\gamma_{1V}(\mathbf{x})$ be the canonical skewness vector of the p -dimensional random vector $\mathbf{x}$. Additionally, let $\gamma_{1V}(\mathbf{A}\mathbf{x} + \mathbf{b})$ be the canonical skewness vector of $\mathbf{A}\mathbf{x} + \mathbf{b}$, where $\mathbf{A}$ is a nonsingular $p \times p$ matrix and $\mathbf{b}$ is a p -dimensional real vector. Finally, let $\mathbf{U} = \boldsymbol{\Sigma}^{-1/2} \mathbf{A}^{-1} \boldsymbol{\Omega}^{1/2}$, where $\boldsymbol{\Sigma}^{-1/2}$ and $\boldsymbol{\Omega}^{1/2}$ are the positive definite symmetric square roots of the concentration matrix of $\mathbf{x}$ and of the covariance matrix of $\mathbf{A}\mathbf{x} + \mathbf{b}$. Then, $\mathbf{U}$ is orthogonal, coincides with $\mathbf{A}$ when $\mathbf{A}$ itself is orthogonal and $\gamma_{1V}(\mathbf{x}) = \mathbf{U}\gamma_{1V}(\mathbf{A}\mathbf{x} + \mathbf{b})$.

Let $X_1, \dots, X_n$ be independent and identically distributed random variables whose third standardized cumulant is γ_1 . It is well-known that the third cumulant of the sum $X_1 + \dots + X_n$ is $\gamma_1/\sqrt{n}$. The following theorem generalizes this statement to multivariate random vectors.

Theorem 4. Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be n independent and identically distributed p -dimensional random vectors with positive definite covariance matrices and finite third-order moments. Additionally, let $\gamma_{1V}(\mathbf{x}_i)$, $\beta_{1P}(\mathbf{x}_i)$ and $\beta_{1T}(\mathbf{x}_i)$ be the canonical skewness vector, the partial skewness and the total skewness of $\mathbf{x}_i$, for $i = 1, \dots, n$. Finally, let $\gamma_{1V}(\mathbf{s}_n)$, $\beta_{1P}(\mathbf{s}_n)$ and $\beta_{1T}(\mathbf{s}_n)$ be the canonical

skewness vector, the partial skewness and the total skewness of $\mathbf{s}_n = \mathbf{x}_1 + \dots + \mathbf{x}_n$. Then, the following identities hold true:

$$\gamma_{1V}(\mathbf{s}_n) = \frac{\gamma_{1V}(\mathbf{x}_i)}{\sqrt{n}}, \beta_{1P}(\mathbf{s}_n) = \frac{\beta_{1P}(\mathbf{x}_i)}{n} \text{ and } \beta_{1T}(\mathbf{s}_n) = \frac{\beta_{1T}(\mathbf{x}_i)}{n}.$$

The total skewness and the partial skewness are the squared norms of the third standardized moment matrix and the canonical skewness vector. Therefore, both skewness measures are functions of the third standardized moments. The following theorem shows that they can also be represented as functions of the second and third central moments, by means of simple matrix operations acting on the covariance and on the third cumulant. Thus, the following theorem provides an alternative representation for the partial and total skewness measures.

Theorem 5. Let $\mathbf{x}$ be a p -dimensional random vector with finite third-order moments and positive definite covariance matrix. Additionally, let $\mathbf{K}_{3,\mathbf{x}}$ and Σ^{-1} be the third cumulant and the concentration matrix of $\mathbf{x}$. Then, the total skewness and the partial skewness of $\mathbf{x}$ are as follows:

$$\beta_{1T} = \text{tr} \left\{ \Sigma^{-1} \mathbf{K}_{3,\mathbf{x}}^{\top} \left(\Sigma^{-1} \otimes \Sigma^{-1} \right) \mathbf{K}_{3,\mathbf{x}} \right\} \text{ and}$$

$$\beta_{1P} = \text{vec}^{\top} \left(\Sigma^{-1} \right) \left(\mathbf{K}_{3,\mathbf{x}} \Sigma^{-1} \mathbf{K}_{3,\mathbf{x}}^{\top} \right) \text{vec} \left(\Sigma^{-1} \right).$$

The fourth-order standardized moment of a p -dimensional random vector $\mathbf{x}$ with finite fourth-order moments and positive definite covariance is as follows:

$$\mathbf{M}_{4,\mathbf{z}} = \mathbb{E} \left(\mathbf{z} \otimes \mathbf{z}^{\top} \otimes \mathbf{z} \otimes \mathbf{z}^{\top} \right),$$

where $\mathbf{z}$ is the standardization of $\mathbf{x}$. Refs. [11,15] defined the partial kurtosis (also known as Mardia's kurtosis) and the total kurtosis (also known as Koziol's kurtosis) as the trace and the squared norm of the following: $\mathbf{M}_{4,\mathbf{z}}$:

$$\beta_{2P}(\mathbf{x}) = \text{tr}(\mathbf{M}_{4,\mathbf{z}}) \text{ and } \beta_{2T}(\mathbf{x}) = \|\mathbf{M}_{4,\mathbf{z}}\|^2.$$

Ref. [1] as well as [16] provided some skewness–kurtosis inequalities involving partial skewness. The following theorem contributes to the literature on inequalities between multivariate measures of skewness and kurtosis.

Theorem 6. Let $\mathbf{x}$ be a p -dimensional random vector with a positive definite covariance matrix and finite fourth-order moments. Additionally, let β_{1T} , β_{1P} , β_{2T} , and β_{2P} be the total skewness, the partial skewness, the total kurtosis and the partial kurtosis of $\mathbf{x}$. Then, the following inequalities hold true:

$$\beta_{1T} \leq \frac{p - p^2 + \beta_{2T}}{2} \text{ and } \beta_{1P} \leq \frac{p - p^4 + \beta_{2P}^2}{2}.$$

As a corollary, the inequality $2\gamma_1^2 \leq \beta_2$ holds true for a random variable X whose third and fourth standardized moments are γ_1 and β_2 .

3. Simulations

In this section, we use simulations for comparing the powers of symmetry tests based on partial and total skewness. To this end, we simulated from the two-component normal mixture with proportional covariances and the multivariate skew-normal distribution. For both models, testing symmetry is a crucial issue, where ordinary likelihood-based methods are somewhat problematic. For each choice of the parameters, variables and units, we computed the percentage of samples rejected at the 0.05 level by the testing procedure

proposed by [1,11]. We refer to the former and the latter testing procedures as “Mori” and “Mardia”, respectively. The results given by the simulations clearly hint that “Mori” is a strong competitor for “Mardia”.

We first simulated 10,000 samples of 100 observations from the mixture $\pi_1 N_p(\mathbf{0}_p, \mathbf{I}_p) + (1 - \pi_1) N_p(\mathbf{1}_p, c \cdot \mathbf{I}_p)$, for $\pi_1 = 0.1, 0.2, 0.3, 0.4$, $p = 5, 10, 15, 20$ and $c = 0.5, 1, 2$. The symbols $\mathbf{0}_p$, $\mathbf{1}_p$ and $\mathbf{I}_p$ denote the p -dimensional vector of zeros, the p -dimensional vector of ones and the p -dimensional identity matrix, respectively. The two-component normal mixture with proportional covariances describes the worst case from a multivariate outlier detection perspective and leads to an unbounded likelihood function (see, for example, ref. [5]). The scatter plot in Figure 1 depicts 10,000 outcomes from $0.2 \cdot N_2(\mathbf{0}_2, \mathbf{I}_2) + 0.8 \cdot N_2(2.5 \cdot \mathbf{1}_2, \mathbf{I}_2)$ and exemplifies the skewness of normal mixtures.

Figure 1. Scatterplot of 10,000 outcomes from $0.2 \cdot N_2(\mathbf{0}_2, \mathbf{I}_2) + 0.8 \cdot N_2(2.5 \cdot \mathbf{1}_2, \mathbf{I}_2)$.

The results of this simulation study are reported in Table 1 and may be summarized as follows. When less weight is placed on the more dispersed component, “Mori” always outperforms “Mardia”, although both tests show a very high power. When the components are equally dispersed, “Mardia” always outperforms “Mori”, although only slightly. Both tests show a very low power, which tends to decrease with the absolute difference between the components’ weights. When more weight is placed on the more dispersed component, “Mori” always outperforms “Mardia”, except when the weight of the less dispersed component is 0.1. The powers of “Mardia” and “Mori” increase as $|2\pi_1 - 1|$ increases, that is, as the absolute difference between the mixture weights π_1 and $1 - \pi_1$ increases.

We also simulated 10,000 samples of sizes $n = 100, 200, 500$ from the multivariate skew-normal density function $2\phi_p(\mathbf{x}; \mathbf{0}_p, \mathbf{I}_p) \Phi(\alpha \mathbf{1}_p^\top \mathbf{x})$, where $\alpha = 1, 2, 3, 4$ and $\phi_p(\mathbf{x}; \mathbf{0}_p, \mathbf{I}_p)$ is the probability density function of a p -dimensional standard normal distribution, with $p = 5, 10, 15$, and 20. The information matrix of the multivariate skew-normal distribution is singular when $\alpha = 0$, thus preventing the use of standard likelihood-based methods when testing symmetry. Table 2 reports the results given in this simulation study, which somewhat differ from those in the previous one. Firstly, “Mardia” nearly always outperforms “Mori”, although only slightly. Secondly, the performance of both “Mardia” and “Mori” improves with the number of units and the parameter α , but deteriorate with the number of variables. Thirdly, the difference between the performance of “Mardia” and “Mori” tends to decrease as the number of units increases. The scatter plot in Figure 2 depicts 10,000 outcomes from $2\phi_2(x_1, x_2; \mathbf{0}_2, \mathbf{I}_2) \Phi(2.5x_1 + 2.5x_2)$ and exemplifies the skewness of skew-normal distributions.

Figure 2. Scatterplot of 10,000 outcomes from $2\phi_2(x_1, x_2; \mathbf{0}_2, \mathbf{I}_2)\Phi(2.5x_1 + 2.5x_2)$.

Table 1. For samples generated from normal mixture distributions, the columns contain the percentages of samples rejected by the test based on Mardia's skewness (header: Mardia) and by the test based on partial skewness (header: Mori) at the 0.05 level for $\pi_1 = 0.1, 0.2, 0.3, 0.4$, $p = 5, 10, 15, 20$ and $c = 0.5, 1, 2$.

p	π_1	$c = 0.5$		$c = 1$		$c = 2$	
		Mardia	Mori	Mardia	Mori	Mardia	Mori
5	0.1	82.29	87.81	10.82	9.82	5.62	4.81
	0.2	86.94	97.32	7.47	7.30	10.19	10.31
	0.3	78.89	96.74	5.41	4.68	18.55	25.88
	0.4	66.24	92.45	3.45	3.08	31.33	48.33
10		Mardia	Mori	Mardia	Mori	Mardia	Mori
	0.1	93.90	97.50	10.83	8.64	6.33	4.08
	0.2	94.88	99.83	6.11	5.23	15.23	19.06
	0.3	92.37	99.92	3.79	3.11	32.18	52.37
15	0.4	85.82	99.68	2.63	2.10	53.60	82.97
		Mardia	Mori	Mardia	Mori	Mardia	Mori
	0.1	95.94	98.56	7.12	4.86	6.19	3.97
	0.2	98.29	99.92	3.81	2.52	20.69	29.58
20	0.3	97.53	99.98	2.45	1.45	47.77	75.36
	0.4	95.74	99.98	1.66	0.89	72.81	95.62
		Mardia	Mori	Mardia	Mori	Mardia	Mori
	0.1	96.13	98.31	4.10	2.08	5.67	3.76
20	0.2	99.08	99.96	1.88	0.91	26.45	38.11
	0.3	99.28	99.99	1.17	0.47	61.38	84.68
	0.4	98.61	99.99	0.78	0.33	86.13	98.41

Table 2. For samples generated from skew-normal distributions, the columns contain the percentages of samples rejected by the test based on Mardia's skewness (header: Mardia) and by the test based on partial skewness (header: Mori) at the 0.05 level for $\alpha = 1, 2, 3, 4$, $p = 5, 10, 15, 20$ and $n = 100, 200, 500$.

p	α	$n = 100$		$n = 200$		$n = 500$	
		Mardia	Mori	Mardia	Mori	Mardia	Mori
5	1	12.58	11.59	24.05	24.08	67.51	62.43
	2	26.67	24.97	62.61	58.06	99.55	97.52
	3	33.46	32.24	75.08	68.85	99.96	99.17
	4	35.99	34.40	78.95	72.82	99.96	99.48
10		Mardia	Mori	Mardia	Mori	Mardia	Mori
	1	6.92	5.81	15.57	15.73	50.83	49.70
	2	10.48	9.15	28.48	28.06	84.04	77.13
	3	11.33	10.10	31.83	31.83	88.89	82.61
15	4	11.82	10.20	32.90	32.66	91.34	85.02
		Mardia	Mori	Mardia	Mori	Mardia	Mori
	1	4.17	2.45	10.06	8.56	33.29	34.03
	2	4.88	3.04	13.55	13.05	51.70	51.34
20	3	4.94	3.19	15.52	14.23	57.02	55.55
	4	5.38	3.22	15.81	15.01	58.27	56.17
		Mardia	Mori	Mardia	Mori	Mardia	Mori
	1	2.19	0.75	6.02	5.32	21.27	23.43
20	2	2.52	1.00	7.69	6.78	30.64	33.01
	3	2.36	1.02	7.82	7.12	32.63	34.13
	4	2.25	0.97	8.39	7.02	33.89	35.27

4. Example

This section uses the financial data in [17] to illustrate a statistical application of the canonical skewness vector and the partial skewness. The dataset, which we shall refer to as ALL, includes the percentage logarithmic daily returns recorded in the French, Dutch and Spanish financial markets from 25 June 2003 to 23 June 2008. Let GOOD be the subset of ALL, which includes the returns in ALL following a positive return of the U.S. financial market, that is, good news from the leading world financial market. Similarly, let BAD be the subset of ALL, which includes all the returns in ALL following a negative return of the U.S. financial market, that is, bad news from the leading world financial market. The three datasets are instrumental in investigating the behavior of financial markets in the presence of either good or bad news [18,19].

Let X_i be the i -th univariate observation in a sample of size n . The sample mean, the sample variance and the sample skewness are as follows:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i, S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 \text{ and } g_1 = \frac{1}{n} \sum_{i=1}^n \left(\frac{X_i - \bar{X}}{S} \right)^3.$$

The sample skewnesses of the French, Danish and Spanish data in ALL are negative: -0.349 , -0.236 and -0.461 . The sample skewnesses of the French, Danish and Spanish data in BAD are negative as well: -0.942 , -0.734 and -1.283 . On the other hand, the sample skewnesses of the French, Danish and Spanish data in GOOD are positive: 0.578 , 0.585 and 0.774 . The absolute skewnesses of the three markets in ALL are smaller than the corresponding absolute skewnesses in GOOD, which in turn are smaller than the corresponding absolute skewnesses in BAD. All these features are well-known stylized facts of financial markets and are modeled by the SGARCH model in [18].

Let the vector $\mathbf{x}_i$ be the i -th multivariate observation in a sample of size n . The sample mean, the sample variance, the sample canonical skewness vector and the sample partial skewness are as follows:

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i, \mathbf{S} = \frac{1}{n} \sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^\top, \mathbf{z}_i = \mathbf{S}^{-1/2}(\mathbf{x}_i - \bar{\mathbf{x}}),$$

$$\mathbf{g}_{1V} = \frac{1}{n} \sum_{i=1}^n \mathbf{z}_i^\top \mathbf{z}_i \mathbf{z}_i \text{ and } b_{1P} = \|\mathbf{g}_{1V}\|^2 = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \mathbf{z}_i^\top \mathbf{z}_i \mathbf{z}_j^\top \mathbf{z}_j \mathbf{z}_i^\top \mathbf{z}_j.$$

In the ALL, GOOD and BAD datasets, $\mathbf{x}_i$ is the trivariate vector whose elements are the returns of the French, Dutch and Spanish financial markets recorded during the i -th day. The canonical skewness vectors of ALL, GOOD and BAD are negative, positive and negative:

$$\mathbf{g}_{1V}^{ALL} = \begin{pmatrix} -0.359 \\ -0.111 \\ -0.341 \end{pmatrix}, \mathbf{g}_{1V}^{GOOD} = \begin{pmatrix} 0.275 \\ 0.362 \\ 0.633 \end{pmatrix}, \mathbf{g}_{1V}^{BAD} = \begin{pmatrix} -0.771 \\ -0.295 \\ -1.008 \end{pmatrix}.$$

Their squared norms, that is, the partial skewnesses of ALL, GOOD and BAD are as follows:

$$b_{1P}^{ALL} = 0.258, b_{1P}^{GOOD} = 0.607, b_{1P}^{BAD} = 1.697.$$

The partial skewness of ALL is smaller than the partial skewness of GOOD, which in turn is smaller than the partial skewness of BAD. Therefore, for the data at hand, the canonical skewness vector and the partial skewness nicely generalize to the multivariate case and the univariate features of financial returns.

5. Conclusions

This paper investigated some theoretical and empirical properties of the canonical skewness vector. The theoretical contribution is threefold. Theorems 1 and 2 pertain to the representation of the canonical skewness vector, either with the star product or with some location vectors. Theorems 3 and 4 deal with transformations, which may be either linear transformations or convolutions. Theorems 5 and 6 highlight some connections between the partial skewness and the total skewness, with respect to representations and inequalities. The empirical contribution is twofold. Firstly, the results of the simulation studies hint that partial skewness might outperform total skewness as a test statistic for symmetry. Secondly, the canonical skewness vector, when applied to the financial data in Section 4, generalizes to multivariate financial returns—a well-known feature of univariate financial returns.

Regrettably, the literature on the analytical forms and properties of the canonical skewness vector under well-known parametric assumptions is quite sparse. For example, to the best of our knowledge, there does not exist any theoretical result regarding the canonical skewness vector of skew-elliptical distributions [20], which are very useful in modeling linear functions of order statistics (see, for example, ref. [21]). We believe that the problem might be conveniently addressed within the smaller class of scale mixtures of skew-normal distributions by jointly using the results on their moments [22–24] and those in this paper (see, for example, Theorem 5). Additionally, the projection of a random vector onto the direction of its canonical skewness vector has a simple parametric interpretation when the underlying distribution is a location mixture of two normal distributions with proportional covariance matrices [5]. It would be interesting to know whether this result might be somewhat extended to scale mixtures of skew-normal distributions.

For the dataset in Section 4, the canonical skewness vector and the partial skewness generalize to the multivariate case the tendency of univariate financial returns to be more skewed in the presence of bad news. This empirical result encourages addressing the

following research questions. Firstly, do the results for the dataset in Section 4 hold for other financial markets and for other time periods? Secondly, could the canonical skewness vector be meaningfully used to measure the effect of news about the COVID-19 pandemic on financial markets? Thirdly, does the canonical skewness vector or the partial skewness have a simple analytical form under the multivariate SGARCH model? Fourthly, is the partial skewness of several small, capitalized financial markets smaller than the partial skewness of the same number of larger financial markets? We hope to address these interesting research questions in future works.

Funding: This research received no external funding.

Data Availability Statement: The dataset used in this paper is available from Morgan Stanley Capital International (<https://www.msci.com/>, accessed on 1 September 2021).

Acknowledgments: The author would like to thank three anonymous reviewers whose insightful comments greatly helped to improve the quality of this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Proof of Theorem 1. By definition, the third standardized cumulant of $\mathbf{x}$ is as follows:

$$\mathbf{K}_{3,\mathbf{z}} = \mathbf{E}(\mathbf{z} \otimes \mathbf{z}^\top \otimes \mathbf{z}),$$

which might be represented as follows:

$$\mathbf{K}_{3,\mathbf{z}} = \mathbf{E}(\mathbf{z} \otimes \mathbf{z}\mathbf{z}^\top)$$

by applying the following identities:

$$\mathbf{a}\mathbf{b}^\top = \mathbf{a} \otimes \mathbf{b}^\top = \mathbf{b}^\top \otimes \mathbf{a},$$

where $\mathbf{a}$ and $\mathbf{b}$ are any two vectors ([25], page 199). As a direct consequence, the third standardized cumulant might be represented as a block column vector:

$$\mathbf{K}_{3,\mathbf{z}} = \begin{pmatrix} \mathbf{K}_1 \\ \dots \\ \mathbf{K}_p \end{pmatrix}, \text{ where } \mathbf{K}_i = \mathbf{E}(Z_i \mathbf{z}\mathbf{z}^\top), \text{ for } i \in \{1, \dots, p\}.$$

Let $\mathbf{k}_{ij}$ be the j -th column of the matrix $\mathbf{K}_i$:

$$\mathbf{k}_{ij} = \mathbf{E}(Z_i Z_j \mathbf{z}), \text{ for } i, j \in \{1, \dots, p\}.$$

By definition, the star product of the $p \times p$ identity matrix $\mathbf{I}_p$ and the third standardized cumulant $\mathbf{K}_{3,\mathbf{z}}$ is the sum of the vectors $\mathbf{k}_{ij}$ multiplied by the elements of $\mathbf{I}_p$ with the same indices, that is, the following:

$$\mathbf{I}_p \star \mathbf{K}_{3,\mathbf{z}} = \sum_{i=1}^p \sum_{j=1}^p \mathbf{k}_{ij} \delta_{ij},$$

where δ_{ij} is the element of $\mathbf{I}_p$ belonging to its i -th row and to its j -th column, given as follows:

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

The definitions of $\mathbf{k}_{ij}$ and δ_{ij} , together with linear properties of the expectation, yield the following:

$$\mathbf{I}_p \star \mathbf{K}_{3,\mathbf{z}} = \sum_{i=1}^p \sum_{j=1}^p \mathbb{E}(Z_i Z_j \mathbf{z}) \delta_{ij} = \sum_{i=1}^p \mathbb{E}(Z_i^2 \mathbf{z}) = \mathbb{E}\left(\sum_{i=1}^p Z_i^2 \mathbf{z}\right).$$

The last expectation coincides with the definition of the canonical skewness vector. The proof is then complete. $\square$

Proof of Theorem 2. By definition, the vectorial skewness of $\mathbf{x}$ is as follows:

$$\gamma_{1V} = \mathbb{E}(\mathbf{z}^\top \mathbf{z} \mathbf{z}), \text{ where } \mathbf{z} = \Sigma^{-1/2}(\mathbf{x} - \mu).$$

The replacement of $\mathbf{z}$ with $\Sigma^{-1/2}(\mathbf{x} - \mu)$ in the definition of γ_{1V} yields the following:

$$\gamma_{1V} = \mathbb{E}\left\{(\mathbf{x} - \mu)^\top \Sigma^{-1}(\mathbf{x} - \mu) \Sigma^{-1/2}(\mathbf{x} - \mu)\right\}.$$

Applying the linear properties of the expectation yields the following:

$$\gamma_{1V} = \mathbb{E}(\mathbf{z}^\top \mathbf{z} \mathbf{z}) = \Sigma^{-1/2} \left\{ \mathbb{E}(\mathbf{z}^\top \mathbf{z} \mathbf{z}) - \mu \mathbb{E}(\mathbf{z}^\top \mathbf{z}) \right\} = \Sigma^{-1/2} \left[\mathbb{E}\left\{(\mathbf{x} - \mu)^\top \Sigma^{-1}(\mathbf{x} - \mu) \mathbf{x}\right\} - p\mu \right]$$

where the last equality follows by taking into account that $\mathbf{z}$ is a random vector having a zero mean vector and a covariance matrix equal to the identity matrix $\mathbf{I}_p$.

We now use two properties of the vectorization and Kronecker operators: (1) $\mathbf{a} \mathbf{b}^\top = \mathbf{a} \otimes \mathbf{b}^\top = \mathbf{b}^\top \otimes \mathbf{a}$ (2) $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^\top \otimes \mathbf{A}) \text{vec}(\mathbf{B})$, where $\mathbf{A} \in \mathbb{R}^p \times \mathbb{R}^q$, $\mathbf{B} \in \mathbb{R}^q \times \mathbb{R}^r$ and $\mathbf{C} \in \mathbb{R}^r \times \mathbb{R}^s$. Applying these properties, we obtain the following:

$$\mathbf{K}_{3,\mathbf{x}} = \mathbb{E}\left[(\mathbf{x} - \mu) \otimes (\mathbf{x} - \mu)^\top \otimes (\mathbf{x} - \mu)\right] = \mathbb{E}\left[(\mathbf{x} - \mu) \otimes (\mathbf{x} - \mu)(\mathbf{x} - \mu)^\top\right]$$

and $\mathbf{K}_{3,\mathbf{x}}^\top = \mathbb{E}\left[(\mathbf{x} - \mu)^\top \otimes (\mathbf{x} - \mu)(\mathbf{x} - \mu)^\top\right]$, from which we obtain the following:

$$\begin{aligned} \mathbf{K}_{3,\mathbf{x}}^\top \text{vec}(\Sigma^{-1}) &= \mathbb{E}\left[\left\{(\mathbf{x} - \mu)^\top \otimes (\mathbf{x} - \mu)(\mathbf{x} - \mu)^\top\right\} \text{vec}(\Sigma^{-1})\right] \\ &= \mathbb{E}\left[\text{vec}\left((\mathbf{x} - \mu)(\mathbf{x} - \mu)^\top \Sigma^{-1}(\mathbf{x} - \mu)\right)\right] = \mathbb{E}\left[(\mathbf{x} - \mu)(\mathbf{x} - \mu)^\top \Sigma^{-1}(\mathbf{x} - \mu)\right] \\ &= \mathbb{E}\left[\mathbf{x}(\mathbf{x} - \mu)^\top \Sigma^{-1}(\mathbf{x} - \mu)\right] - \mathbb{E}\left[(\mathbf{x} - \mu)^\top \Sigma^{-1}(\mathbf{x} - \mu)\right] \mu = \mathbb{E}\left[\mathbf{x}(\mathbf{x} - \mu)^\top \Sigma^{-1}(\mathbf{x} - \mu)\right] - p\mu \end{aligned}$$

from which the following holds true:

$$\Sigma^{-1/2} \mathbf{K}_{3,\mathbf{x}}^\top \text{vec}(\Sigma^{-1}) = \Sigma^{-1/2} \left[\mathbb{E}\left\{\mathbf{x}(\mathbf{x} - \mu)^\top \Sigma^{-1}(\mathbf{x} - \mu)\right\} - p\mu \right] = \gamma_{1V},$$

as we aimed to prove. $\square$

Proof of Theorem 3. Let Σ , Σ^{-1} and $\Sigma^{-1/2}$ be the covariance matrix of $\mathbf{x}$, the concentration matrix of $\mathbf{x}$ and the positive definite symmetric square root of Σ^{-1} . Additionally, let Ω , Ω^{-1} and $\Omega^{-1/2}$ be the covariance matrix of $\mathbf{Ax} + \mathbf{b}$, the concentration matrix of $\mathbf{Ax} + \mathbf{b}$ and the positive definite symmetric square root of Ω^{-1} . Let us also denote by $\mathbf{A}^{-\top}$ the transpose of the matrix $\mathbf{A}^{-1}$.

We first prove that $\mathbf{U} = \Sigma^{-1/2} \mathbf{A}^{-1} \Omega^{1/2}$ is orthogonal as follows:

$$\mathbf{U}\mathbf{U}^\top = \mathbf{U}^\top \mathbf{U} = \left(\Sigma^{-1/2} \mathbf{A}^{-1} \Omega^{1/2} \right) \left(\Sigma^{-1/2} \mathbf{A}^{-1} \Omega^{1/2} \right)^\top =$$

$$\Sigma^{-1/2} \mathbf{A}^{-1} \Omega^{1/2} \Omega^{1/2} \mathbf{A}^{-\top} \Sigma^{-1/2} = \Sigma^{-1/2} \mathbf{A}^{-1} \Omega \mathbf{A}^{-\top} \Sigma^{-1/2}.$$

The definitions of Ω and $\Sigma^{-1/2}$ imply the following identities:

$$\Omega = \mathbf{A} \Sigma \mathbf{A}^\top \text{ and } \Sigma^{-1/2} \Sigma \Sigma^{-1/2} = \mathbf{I}_p,$$

so that the product of $\mathbf{U}$ and its transpose is an identity matrix as follows:

$$\mathbf{U}\mathbf{U}^\top = \Sigma^{-1/2} \mathbf{A}^{-1} \Omega \mathbf{A}^{-\top} \Sigma^{-1/2} = \Sigma^{-1/2} \mathbf{A}^{-1} \mathbf{A} \Sigma \mathbf{A}^\top \mathbf{A}^{-\top} \Sigma^{-1/2} = \Sigma^{-1/2} \Sigma \Sigma^{-1/2} = \mathbf{I}_p.$$

Therefore $\mathbf{U}$ is an orthogonal matrix; this concludes the first part of the proof.

We assume without loss of generality that $\mathbf{x}$ is centered at the origin and that $\mathbf{b}$ is a null vector: $E(\mathbf{x}) = \mathbf{b} = \mathbf{0}_p$. The following identities hold true when $\mathbf{A}$ is an orthogonal matrix:

$$\mathbf{A}^{-1} = \mathbf{A}^\top, \Omega^{-1/2} = \mathbf{A} \Sigma^{-1/2} \mathbf{A}^\top.$$

The canonical skewness vector of $\mathbf{A}\mathbf{x}$ is then the following:

$$\gamma_{1V}(\mathbf{A}\mathbf{x}) = E\left(\mathbf{x}^\top \Sigma^{-1} \mathbf{x} \Omega^{-1/2} \mathbf{A}\mathbf{x}\right) = E\left(\mathbf{x}^\top \Sigma^{-1} \mathbf{x} \mathbf{A} \Sigma^{-1/2} \mathbf{A}^\top \mathbf{A}\mathbf{x}\right).$$

Applying now the linear properties of the expected value completes the second part of the proof:

$$\gamma_{1V}(\mathbf{A}\mathbf{x}) = E\left(\mathbf{x}^\top \Sigma^{-1} \mathbf{x} \mathbf{A} \Sigma^{-1/2} \mathbf{x}\right) = \mathbf{A} E\left(\mathbf{x}^\top \Sigma^{-1} \mathbf{x} \Sigma^{-1/2} \mathbf{x}\right) = \mathbf{A} \gamma_{1V}(\mathbf{x}).$$

By definition, the canonical skewness vectors of $\mathbf{x}$ and $\mathbf{A}\mathbf{x}$ are as follows:

$$\gamma_{1V}(\mathbf{x}) = E\left(\mathbf{x}^\top \Sigma^{-1} \mathbf{x} \Sigma^{-1/2} \mathbf{x}\right) \text{ and}$$

$$\gamma_{1V}(\mathbf{A}\mathbf{x}) = E\left[(\mathbf{A}\mathbf{x})^\top \Omega^{-1} (\mathbf{A}\mathbf{x}) \Omega^{-1/2} \mathbf{A}\mathbf{x}\right],$$

By the ordinary properties of matrix inverses and the definition of Ω , we have the following:

$$\Omega^{-1} = \mathbf{A}^{-\top} \Sigma^{-1} \mathbf{A}^{-1}.$$

The canonical skewness vector of $\mathbf{A}\mathbf{x}$ is then the following:

$$\gamma_{1V}(\mathbf{A}\mathbf{x}) = E\left(\mathbf{x}^\top \Sigma^{-1} \mathbf{x} \Omega^{-1/2} \mathbf{A}\mathbf{x}\right).$$

Apply now the linear properties of the expected value and recall the definition of $\mathbf{U}$ to obtain the following identity:

$$\gamma_{1V}(\mathbf{x}) = \Sigma^{-1/2} \mathbf{A}^{-1} \Omega^{1/2} \gamma_{1V}(\mathbf{A}\mathbf{x}) = \mathbf{U} \gamma_{1V}(\mathbf{A}\mathbf{x}).$$

□

Proof of Theorem 4. The canonical skewness vector, the partial skewness, and the total skewness are location invariant. Therefore, we can assume without loss of generality that the means of $\mathbf{x}_i, \dots, \mathbf{x}_n$ and of their sum $\mathbf{s}_n$ coincide with the p -dimensional null vector as follows:

$$E(\mathbf{x}_i) = E(\mathbf{s}_n) = \mathbf{0}_p.$$

Let Σ , Σ^{-1} and $\Sigma^{-1/2}$ be the covariance matrix of $\mathbf{x}_i$, the concentration matrix of $\mathbf{x}_i$ and the positive definite symmetric square root of Σ^{-1} . Then, the covariance matrix of $\mathbf{s}_n$, the concentration matrix of $\mathbf{s}_n$ and its positive definite symmetric square root are as follows:

$$n\Sigma, \frac{\Sigma^{-1}}{n} \text{ and } \frac{\Sigma^{-1/2}}{\sqrt{n}}.$$

By definition, the canonical skewness vectors of $\mathbf{x}_i$ and $\mathbf{s}_n$ are as follows:

$$\gamma_{1V}(\mathbf{x}_i) = E\left(\mathbf{x}_i^\top \Sigma^{-1} \mathbf{x}_i \Sigma^{-1/2} \mathbf{x}_i\right) \text{ and } \gamma_{1V}(\mathbf{s}_n) = E\left(\mathbf{s}_n^\top \frac{1}{n} \Sigma^{-1} \mathbf{s}_n \frac{1}{\sqrt{n}} \Sigma^{-1/2} \mathbf{s}_n\right).$$

The definitions of $\mathbf{s}_n$ and the linear properties of the expectation yield the following identities:

$$\begin{aligned} \gamma_{1V}(\mathbf{s}_n) &= E\left(\mathbf{s}_n^\top \frac{1}{n} \Sigma^{-1} \mathbf{s}_n \frac{1}{\sqrt{n}} \Sigma^{-1/2} \mathbf{s}_n\right) = \frac{1}{\sqrt{n^3}} \Sigma^{-1/2} E\left(\mathbf{s}_n^\top \Sigma^{-1} \mathbf{s}_n \mathbf{s}_n\right) = \\ &= \frac{1}{\sqrt{n^3}} \Sigma^{-1/2} \sum_{i=1}^n E\left(\mathbf{s}_n^\top \Sigma^{-1} \mathbf{s}_n \mathbf{x}_i\right) = \frac{1}{\sqrt{n^3}} \Sigma^{-1/2} \sum_{i=1}^n E\left(\sum_{j=1}^n \sum_{h=1}^n \mathbf{x}_j^\top \Sigma^{-1} \mathbf{x}_h \mathbf{x}_i\right). \end{aligned}$$

By assumption, the random vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ are mutually independent and centered. Therefore, the mean of $\mathbf{x}_j^\top \Sigma^{-1} \mathbf{x}_h \mathbf{x}_i$ is the p -dimensional null vector if at least one of the three indices differs from the other two:

$$E\left(\sum_{j=1}^n \sum_{h=1}^n \mathbf{x}_j^\top \Sigma^{-1} \mathbf{x}_h \mathbf{x}_i\right) = \mathbf{0}_p \text{ if either } i \neq j, i \neq h \text{ or } j \neq h.$$

The above identity and linear properties of the expectation yield the following:

$$\gamma_{1V}(\mathbf{s}_n) = \frac{1}{\sqrt{n^3}} \Sigma^{-1/2} \sum_{i=1}^n E\left(\mathbf{x}_i^\top \Sigma^{-1} \mathbf{x}_i \mathbf{x}_i\right) = \frac{1}{\sqrt{n^3}} \sum_{i=1}^n E\left(\mathbf{x}_i^\top \Sigma^{-1} \mathbf{x}_i \Sigma^{-1/2} \mathbf{x}_i\right).$$

The random vectors $\mathbf{x}_1, \dots, \mathbf{x}_n$ are identically distributed, and their canonical skewness vector is $\gamma_{1V}(\mathbf{x}_i)$. Therefore, we can complete the first part of the proof:

$$\gamma_{1V}(\mathbf{s}_n) = \frac{1}{\sqrt{n^3}} n \gamma_{1V}(\mathbf{x}_i) = \frac{1}{\sqrt{n}} \gamma_{1V}(\mathbf{x}_i).$$

By definition, the partial skewness of a random vector is just the squared norm of its canonical skewness vector, so we can complete the second part of the proof:

$$\beta_{1P}(\mathbf{s}_n) = \|\gamma_{1V}(\mathbf{s}_n)\|^2 = \gamma_{1V}^\top(\mathbf{s}_n) \gamma_{1V}(\mathbf{s}_n) = \frac{\gamma_{1V}^\top(\mathbf{x}_i) \gamma_{1V}(\mathbf{x}_i)}{n} = \frac{\beta_{1P}(\mathbf{x}_i)}{n}.$$

Let $\mathbf{y}_1, \dots, \mathbf{y}_n$ be n random vectors independent of each other and independent of $\mathbf{x}_1, \dots, \mathbf{x}_n$. Additionally, let $\mathbf{y}_i$ and $\mathbf{x}_j$ be identically distributed, for $i, j = 1, \dots, n$. Finally, let the following hold:

$$\mathbf{t}_n = \sum_{i=1}^n \mathbf{y}_i \text{ and } b_{ij} = \mathbf{x}_i^\top \Sigma^{-1} \mathbf{y}_j.$$

Then, the total skewness of $\mathbf{s}_n$ is as follows:

$$\beta_{1T}(\mathbf{s}_n) = E\left\{\left(\mathbf{s}_n^\top \frac{1}{n} \Sigma^{-1} \mathbf{t}_n\right)^3\right\} = \frac{1}{n^3} E\left\{\left(\mathbf{s}_n^\top \Sigma^{-1} \mathbf{t}_n\right)^3\right\}.$$

By recalling the definitions of $\mathbf{s}_n$, $\mathbf{t}_n$ and $\mathbf{b}_{ij}$, we obtain the following:

$$\beta_{1T}(\mathbf{s}_n) = \frac{1}{n^3} \mathbb{E} \left\{ \left(\sum_{i,j} b_{ij} \right)^3 \right\} = \frac{1}{n^3} \mathbb{E} \left\{ \sum_{i,j} b_{ij}^3 + 3 \sum_{\{i,j\} \neq \{h,k\}} b_{ij}^2 b_{hk} \right\}.$$

By assumption, the random vectors $\mathbf{y}_1, \dots, \mathbf{y}_n, \mathbf{x}_1, \dots, \mathbf{x}_n$ are mutually independent and centered, so the mean of $b_{ij}^2 b_{hk}$ is zero whenever at least one index differs from the others:

$$\mathbb{E}(b_{ij}^2 b_{hk}) = 0 \text{ if either } i \neq j, i \neq h, i \neq k, j \neq h, j \neq k \text{ or } h \neq k.$$

The total skewness of $\mathbf{s}_n$ is then as follows:

$$\beta_{1T}(\mathbf{s}_n) = \frac{1}{n^3} \mathbb{E} \left(\sum_{i,j} b_{ij}^3 \right) = \sum_{i,j} \mathbb{E}(b_{ij}^3).$$

By assumption, the random vectors $\mathbf{y}_i$ and $\mathbf{x}_j$ are identically distributed, for $i, j = 1, \dots, n$:

$$\beta_{1T}(\mathbf{s}_n) = \frac{1}{n^3} \mathbb{E} \left(\sum_{i,j} b_{ij}^3 \right) = \frac{\mathbb{E}(b_{ij}^3)}{n}.$$

We conclude the proof by recalling the definition of the total skewness of $\mathbf{x}_i$:

$$\beta_{1T}(\mathbf{s}_n) = \frac{\beta_{1T}(\mathbf{x}_i)}{n}.$$

□

Proof of Theorem 5. Let $\mathbf{z} = \Sigma^{-1/2}(\mathbf{x} - \boldsymbol{\mu})$ be the standardization of $\mathbf{x}$, where $\boldsymbol{\mu}$ is the mean of $\mathbf{x}$ and $\Sigma^{-1/2}$ is the symmetric, positive definite square root of Σ^{-1} :

$$\Sigma^{-1/2} \Sigma^{-1/2} = \Sigma^{-1}, \Sigma^{-1/2} = \left(\Sigma^{-1/2} \right)^\top, \Sigma^{-1/2} > 0.$$

By definition, the third standardized cumulant of $\mathbf{x}$ is the third cumulant of $\mathbf{z}$, that is, the following:

$$\mathbf{K}_{3,\mathbf{z}} = \mathbb{E}(\mathbf{z} \otimes \mathbf{z}^\top \otimes \mathbf{z}).$$

The total skewness of $\mathbf{x}$ might be represented as the squared Euclidean norm of $\mathbf{K}_{3,\mathbf{z}}$ [26]:

$$\beta_{1T} = \|\mathbf{K}_{3,\mathbf{z}}\|^2 = \text{tr}(\mathbf{K}_{3,\mathbf{z}}^\top \mathbf{K}_{3,\mathbf{z}}).$$

The third standardized cumulant of $\mathbf{x}$ might be represented by means of $\Sigma^{-1/2}$ and $\mathbf{K}_{3,\mathbf{x}}$:

$$\mathbf{K}_{3,\mathbf{z}} = \left(\Sigma^{-1/2} \otimes \Sigma^{-1/2} \right) \mathbf{K}_{3,\mathbf{x}} \Sigma^{-1/2}.$$

The two identities above imply the following one:

$$\beta_{1T} = \text{tr} \left\{ \Sigma^{-1/2} \mathbf{K}_{3,\mathbf{x}}^\top \left(\Sigma^{-1/2} \otimes \Sigma^{-1/2} \right) \left(\Sigma^{-1/2} \otimes \Sigma^{-1/2} \right) \mathbf{K}_{3,\mathbf{x}} \Sigma^{-1/2} \right\}.$$

The trace of the product of two symmetric matrices with the same number of rows does not depend on the order of the matrices themselves:

$$\beta_{1T} = \text{tr} \left\{ \Sigma^{-1/2} \Sigma^{-1/2} \mathbf{K}_{3,\mathbf{x}}^\top \left(\Sigma^{-1/2} \otimes \Sigma^{-1/2} \right) \left(\Sigma^{-1/2} \otimes \Sigma^{-1/2} \right) \mathbf{K}_{3,\mathbf{x}} \right\}.$$

By the distributive property of the Kronecker product and the ordinary matrix product, if matrices **A**, **B**, **C** and **D** are of appropriate size, then the following holds:

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD}$$

(see, for example, ref. [25], page 194). By applying this property to the above identity, we obtain the following:

$$\beta_{1T} = \text{tr} \left\{ \boldsymbol{\Sigma}^{-1/2} \boldsymbol{\Sigma}^{-1/2} \mathbf{K}_{3,x}^\top \left(\boldsymbol{\Sigma}^{-1/2} \boldsymbol{\Sigma}^{-1/2} \otimes \boldsymbol{\Sigma}^{-1/2} \boldsymbol{\Sigma}^{-1/2} \right) \mathbf{K}_{3,x} \right\}.$$

By definition, the product of $\boldsymbol{\Sigma}^{-1/2}$ by itself coincides with the concentration matrix $\boldsymbol{\Sigma}^{-1}$ so that the following holds:

$$\beta_{1T} = \text{tr} \left\{ \boldsymbol{\Sigma}^{-1} \mathbf{K}_{3,x}^\top \left(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} \right) \mathbf{K}_{3,x} \right\}$$

and the first part of the proof is complete.

We now prove the second part of the theorem. As a straightforward implication of Theorem 2, we have the following:

$$\beta_{1P} = \gamma_{1V}^\top \gamma_{1V} = \text{vec} \left(\boldsymbol{\Sigma}^{-1} \right) \mathbf{K}_{3,x} \boldsymbol{\Sigma}^{-1/2} \boldsymbol{\Sigma}^{-1/2} \mathbf{K}_{3,x}^\top \text{vec} \left(\boldsymbol{\Sigma}^{-1} \right) = \text{vec} \left(\boldsymbol{\Sigma}^{-1} \right) \mathbf{K}_{3,x} \boldsymbol{\Sigma}^{-1} \mathbf{K}_{3,x}^\top \text{vec} \left(\boldsymbol{\Sigma}^{-1} \right),$$

as we aimed to prove. $\square$

Proof of Theorem 6. Without loss of generality, we can assume that $\mathbf{x} = (X_1, \dots, X_p)^\top$ is a standard random vector:

$$E(X_i) = 0, E(X_i^2) = 1, E(X_i X_j) = 0, \text{ for } i, j = 1, \dots, p \text{ and } i \neq j.$$

Let $\mathbf{y} = (Y_1, \dots, Y_p)^\top$ be a random vector independent of $\mathbf{x}$ and with its same distribution. The first and second moments of $\mathbf{x}^\top \mathbf{y}$ are as follows:

$$E(\mathbf{x}^\top \mathbf{y}) = E \left(\sum_{i=1}^p X_i Y_i \right) = \sum_{i=1}^p E(X_i Y_i) = \sum_{i=1}^p E(X_i) E(Y_i) = 0 \text{ and}$$

$$E \left\{ (\mathbf{x}^\top \mathbf{y})^2 \right\} = E \left\{ \left(\sum_{i=1}^p X_i Y_i \right)^2 \right\} = \sum_{i=1}^p \sum_{j=1}^p E(X_i Y_i X_j Y_j) = \sum_{i=1}^p E(X_i^2) E(Y_i^2) = p.$$

The variance of the random variable $(\mathbf{x}^\top \mathbf{y})^2 - \mathbf{x}^\top \mathbf{y}$ is nonnegative, thus implying the following inequality:

$$E \left[\left\{ (\mathbf{x}^\top \mathbf{y})^2 - \mathbf{x}^\top \mathbf{y} \right\}^2 \right] \geq p^2.$$

By expanding the square and using linear properties of the expected value, we obtain the following:

$$E \left\{ (\mathbf{x}^\top \mathbf{y})^4 \right\} - 2E \left\{ (\mathbf{x}^\top \mathbf{y})^3 \right\} + E \left\{ (\mathbf{x}^\top \mathbf{y})^2 \right\} \geq p^2.$$

The second, third and fourth moments of $\mathbf{x}^\top \mathbf{y}$ are p , β_{1T} and β_{2T} :

$$\beta_{2T} - 2\beta_{1T} + p \geq p^2 \text{ so that } \beta_{1T} \leq \frac{p - p^2 + \beta_{2T}}{2}.$$

We now prove the second part of the theorem. The expectation of $\mathbf{x}^\top \mathbf{x}$ is as follows:

$$E(\mathbf{x}^\top \mathbf{x}) = E\left(\sum_{i=1}^p X_i^2\right) = \sum_{i=1}^p E(X_i^2) = p.$$

The random vectors $\mathbf{x}$ and $\mathbf{y}$ are independent and identically distributed by the assumption so that the expectation of $(\mathbf{x}^\top \mathbf{x})(\mathbf{y}^\top \mathbf{y}) - \mathbf{x}^\top \mathbf{y}$ is as follows:

$$E\{(\mathbf{x}^\top \mathbf{x})(\mathbf{y}^\top \mathbf{y}) - \mathbf{x}^\top \mathbf{y}\} = p^2.$$

The variance of the random variable $(\mathbf{x}^\top \mathbf{x})(\mathbf{y}^\top \mathbf{y}) - \mathbf{x}^\top \mathbf{y}$ is nonnegative, thus implying the following inequality:

$$E\left\{\left[(\mathbf{x}^\top \mathbf{x})(\mathbf{y}^\top \mathbf{y}) - \mathbf{x}^\top \mathbf{y}\right]^2\right\} \geq p^4.$$

Expansion of the square and linear properties of the expected value yields the following:

$$E\left\{(\mathbf{x}^\top \mathbf{x})^2(\mathbf{y}^\top \mathbf{y})^2\right\} - 2E\{(\mathbf{x}^\top \mathbf{y})(\mathbf{x}^\top \mathbf{x})(\mathbf{y}^\top \mathbf{y})\} + E\{(\mathbf{x}^\top \mathbf{y})^2\} \geq p^4.$$

The partial skewness and the partial kurtosis of $\mathbf{x}$ may be represented as follows (Ref. [27], page 82; ref. [26]):

$$\beta_{1P} = E\{(\mathbf{x}^\top \mathbf{y})(\mathbf{x}^\top \mathbf{x})(\mathbf{y}^\top \mathbf{y})\} \text{ and } \beta_{2P} = E\{(\mathbf{x}^\top \mathbf{x})^2\}$$

thus leading to the following inequality:

$$\beta_{2P}^2 - 2\beta_{1P} + p \geq p^4, \text{ which is equivalent to } \beta_{1P} \leq \frac{p - p^4 + \beta_{2P}^2}{2}.$$

□

References

1. Móri, T.; Rohatgi, V.; Székely, G. On multivariate skewness and kurtosis. *Theory Probab. Its Appl.* **1993**, *38*, 547–551. [\[CrossRef\]](#)
2. Balakrishnan, N.; Brito, M.; Quiroz, A. A Vectorial Notion of Skewness and Its Use in Testing for Multivariate Symmetry. *Commun. Stat.-Theory Methods* **2007**, *36*, 1757–1767. [\[CrossRef\]](#)
3. Kollo, T. Multivariate skewness and kurtosis measures with an application in ICA. *J. Multivar. Anal.* **2008**, *99*, 2328–2338. [\[CrossRef\]](#)
4. Serfling, R.J. Multivariate symmetry and asymmetry. In *Encyclopedia of Statistical Sciences*, 2nd ed.; Kotz, S., Read, C.B., Balakrishnan, N., Vidakovic, B., Eds.; Wiley: New York, NY, USA, 2006.
5. Loperfido, N. Vector-Valued Skewness for Model-Based Clustering. *Stat. Probab. Lett.* **2015**, *99*, 230–237. [\[CrossRef\]](#)
6. Loperfido, N. Singular Value Decomposition of the Third Multivariate Moment. *Linear Algebra Its Appl.* **2015**, *473*, 202–216. [\[CrossRef\]](#)
7. Nordhausen, K.; Oja, H.; Ollila, E. Multivariate models and the first four moments. In *Nonparametric Statistics and Mixture Models—A Festschrift for Thomas P. Hettmansperger*; Hunter, D., Richards, D., Rosenberger, J.L., Eds.; World Scientific Publishing Co Pte Ltd.: Singapore, 2011.
8. Ilmonen, P.; Oja, H.; Serfling, R. On Invariant Coordinate System (ICS) Functionals. *Int. Stat. Rev.* **2012**, *80*, 93–110. [\[CrossRef\]](#)
9. Davis, A. On the Effects of Moderate Multivariate Nonnormality on Wilks's Likelihood Ratio Criterion. *Biometrika* **1980**, *67*, 419–427. [\[CrossRef\]](#)
10. Henze, N. Limit laws for multivariate skewness in the sense of Móri, Rohatgi and Székely. *Stat. Probab. Lett.* **1997**, *33*, 299–307. [\[CrossRef\]](#)
11. Mardia, K. Measures of multivariate skewness and kurtosis with applications. *Biometrika* **1970**, *57*, 519–530. [\[CrossRef\]](#)
12. Malkovich, J.; Afifi, A. On Tests for Multivariate Normality. *J. Am. Stat. Assoc.* **1973**, *68*, 176–179. [\[CrossRef\]](#)
13. MacRae, E. Matrix derivatives with an application to an adaptive linear decision problem. *Ann. Stat.* **1974**, *2*, 337–346. [\[CrossRef\]](#)
14. Kankainen, A.; Taskinen, S.; Oja, H. Tests of multinormality based on location vectors and scatter matrices. *Stat. Methods Appl.* **2007**, *16*, 357–379. [\[CrossRef\]](#)

15. Koziol, J. A note on measures of multivariate kurtosis. *Biom. J.* **1989**, *31*, 619–624. [\[CrossRef\]](#)
16. Ogasawara, H. Extensions of Pearson's inequality between skewness and kurtosis to multivariate cases. *Stat. Probab. Lett.* **2017**, *30*, 12–16. [\[CrossRef\]](#)
17. De Luca, G.; Loperfido, N. Modelling Multivariate Skewness in Financial Returns: A SGARCH Approach. *Eur. J. Financ.* **2015**, *21*, 1113–1131. [\[CrossRef\]](#)
18. De Luca, G.; Loperfido, N. A Skew-in-Mean GARCH Model for Financial Returns. In *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*; Genton, M.G., Ed.; Chapman & Hall, CRC: Boca Raton, FL, USA, 2004; pp. 205–222.
19. De Luca, G.; Genton, M.; Loperfido, N. A Multivariate Skew-Garch Model. In *Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series, Part A (Special Volume in Honor of Robert Engle and Clive Granger, the 2003 Winners of the Nobel Prize in Economics)*; Terrell, D., Ed.; Elsevier: Oxford, UK, 2006; Volume 20, pp. 33–56.
20. Branco, M.; Dey, D. A general class of skew-elliptical distributions. *J. Multivar. Anal.* **2001**, *79*, 99–113. [\[CrossRef\]](#)
21. Loperfido, N. A Note on Skew-Elliptical Distributions and Linear Functions of Order Statistics. *Stat. Probab. Lett.* **2008**, *78*, 3184–3186. [\[CrossRef\]](#)
22. Kim, H.M. Moments of variogram estimator for a generalized skew t distribution. *J. Korean Stat. Soc.* **2005**, *34*, 109–123.
23. Kim, H.M. A note on scale mixtures of skew-normal distributions. *Stat. Probab. Lett.* **2008**, *78*, 1694–1701. [\[CrossRef\]](#)
24. Kim, H.M. Corrigendum to “A note on scale mixtures of skew normal distribution” [*Statist. Probab. Lett.* 78 (2008) 1694–1701]. *Stat. Probab. Lett.* **2013**, *83*, 1937. [\[CrossRef\]](#)
25. Rao, C.; Rao, M. *Matrix Algebra and Its Applications to Statistics and Econometrics*; World Scientific Co. Pte. Ltd.: Singapore, 1998.
26. Kollo, T.; Srivastava, M. Estimation and testing of parameters in multivariate Laplace distribution. *Commun. Stat.-Theory Methods* **2005**, *33*, 2363–2687. [\[CrossRef\]](#)
27. Kotz, S.; Balakrishnan, N.; Johnson, N. *Continuous Multivariate Distributions, Volume 1: Models and Applications*, 2nd ed.; Wiley: New York, NY, USA, 2000.