Singular value decomposition of the third multivariate moment

Nicola Loperfido
Dipartimento di Economia, Società e Politica
Università degli Studi di Urbino "Carlo Bo"
Via Saffi 42, 61029 Urbino, ITALY
email: nicola.loperfido@uniurb.it

Abstract

The third moment of a random vector is a matrix which conveniently arranges all moments of order three which can be obtained from the random vector itself. We investigate some properties of its singular value decomposition. In particular, we show that left eigenvectors corresponding to positive singular values of the third moment are vectorized, symmetric matrices. We derive further properties under the additional assumptions of exchangeability, reversibility and independence. Statistical applications deal with measures of multivariate skewness.

Keywords: Exchangeable process, multivariate skewness, linear process, reversible process, singular value decomposition.

MSC classification: 15A18, 60G09, 60G10, 60G51.

1 Introduction

Let $x = (X_1, ..., X_d)^T$ be a $d-$dimensional random vector satisfying $E\left(\left|X_i^3\right|\right) < +\infty$, for $i = 1, ..., d$. The third moment of x is the $d^2 \times d$ matrix $M_3(x) = E\left(x \otimes x^T \otimes x\right)$, where "$\otimes$" denotes the Kronecker product (see, for example, De Luca and Loperfido, 2012). In the following, when referring to the third moment of a random vector, we shall implicitly assume that all appropriate moments exist. The third central moment of x, also known as its third cumulant, is the third moment of $x - \mu$, where μ is the mean of x. Statistical applications of the third moment include, but are not limited to: factor analysis (Mooijaart, 1985), density approximations (Van Hulle, 2005), Independent Component Analysis (De Lathauwer *et al*, 2001), financial econometrics (De Luca and Loperfido, 2012), cluster analysis (Loperfido, 2013).

No one of the above authors studied the singular value decomposition (SVD, henceforth) of the third moment, which is a fundamental tool in both mathematics and statistics. The role of the SVD in mathematics is well reviewed by

Martin and Porter (2011). In Statistics, the SVD provides the theoretical foundations for the biplot (Gower, 2004), canonical correlation analysis (Mardia *et al*, 1979, page 283) and correspondence analysis (Greenacre and Hastie, 1987).

This paper examines the SVD of the third multivariate moment both in the general case and under additional assumptions. In the general case, it shows that left eigenvectors corresponding to positive singular values of the third moment are vectorized, symmetric matrices. Additional properties are derived for finite realizations of well-known stochastic processes, which nicely mirror those of their second moments. Finally, the paper shows that the SVD is instrumental in obtaining properties of main measures of multivariate skewness.

The rest of the paper is organized as follows. Section 2 discusses the SVD for the third moment, in the general case. Section 3 obtains some inequalities for measures of multivariate skewness. Sections 4 and 5 deal with third moments under independence and invariance assumptions, respectively. Section 6 shows that theorems and proofs similar to those in Section 2 also hold for fourth moments and cumulants. Section 7 applies results in section 3 to financial data.

2 Decomposition

This section investigates the SVD and a related decomposition of the third multivariate moment. The symbols I_d, 1_d and 0_d will denote the $d \times d$ identity matrix, the $d-$dimensional vector of ones and the $d-$dimensional vector of zeros, respectively. Also, $vecA = vec(A)$ will denote the vectorization operator, which stacks the columns of the matrix A on top of one another (Rao and Rao, 1998, page 200). The row vector $vec^T A$ will denote the transpose of the vectorized matrix A, while the column vector $vecA^T$ will denote the vectorized transpose of A.

We shall now recall some fundamental properties of the Kronecker product which we shall use repeatedly in the following proofs (see, for example, Rao and Rao, pages 194-201): (P1) the Kronecker product is associative: $(A \otimes B) \otimes C = A \otimes (B \otimes C) = A \otimes B \otimes C$; (P2) if matrices A, B, C and D are of appropriate size, then $(A \otimes B)(C \otimes D) = AC \otimes BD$ and $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ if both A and B are invertible matrices; (P3) If a and b are two vectors, then ab^T, $a \otimes b^T$ and $b^T \otimes a$ denote the same matrix; (P4) for any two $m \times n$ matrices A and B it holds true that $tr\left(A^T B\right) = vec^T(B)vec(A)$ and when $A = B$, we have $tr(A^T A) = \sigma_1^2 + ... + \sigma_d^2$ with σ_1, ..., σ_d being the singular values of A; (P5) $vec\,(ABC) = \left(C^T \otimes A\right) vec\,(B)$, when $A \in \mathbb{R}^p \times \mathbb{R}^q$, $B \in \mathbb{R}^q \times \mathbb{R}^r$, $C \in \mathbb{R}^r \times \mathbb{R}^s$; (P6) $tr(A \otimes B) = tr(A)tr(B)$, with A and B being two $m \times m$ matrices. Finally, we shall recall some properties of the commutation matrix $K_{p,q} \in \mathbb{R}^{pq} \times \mathbb{R}^{pq}$ (Magnus and Neudecker, 1979): (P7) $vecM^T = K_{p,q}vecM$ for any $p \times q$ matrix M; (P8) $K_{p,q}^T = K_{q,p}$; (P9) $K_{p,p} = K_{p,p}^{-1}$.

Eigenvectors associated with positive eigenvalues of the fourth moment matrix are vectorized, symmetric matrices (Loperfido, 2011). The following theorem shows that a similar result holds for the third multivariate moment.

Theorem 1 *Left singular vectors corresponding to positive singular values of the third multivariate moment are vectorized, symmetric matrices.*

Proof. Let $M_3 = E\left(x \otimes x^T \otimes x\right)$ be the third moment of a $d-$dimensional random vector x, and let $K_{p,q}$ be a $pq \times pq$ commutation matrix. The matrix $x \otimes x^T \otimes x$ can be represented as $vec\left(xx^T\right)x^T$, using standard properties of the Kronecker product and the vectorization operator, so that $M_3 = E\left[vec\left(xx^T\right) \otimes x^T\right]$. Since xx^T is a symmetric matrix, by property P7 we have $vec\left(xx^T\right) = K_{d,d}vec\left(xx^T\right)$ and $M_3 = E\left[K_{d,d}vec\left(xx^T\right) \otimes x^T\right]$. Apply now linear properties of the expected value to obtain $M_3 = K_{d,d}E\left[vec\left(xx^T\right) \otimes x^T\right] = K_{d,d}M_3$. Let σ be a positive singular value corresponding to the left singular vector l of M_3: $Ll = \sigma l$, where $L = M_3 M_3^T$. Finally, let l be the vectorized, $d \times d$ matrix A. Since $M_3 = K_{d,d}M_3$ we have $L = K_{d,d}M_3 M_3^T$ and $K_{d,d}Ll = \sigma l$. By properties P8 and P9, the matrix $K_{d,d}$ is at the same time symmetric and orthogonal, so that $Ll = \sigma K_{d,d} l$. The above equations lead to the following one: $l = K_{d,d}l$, which can be represented as $vecA = K_{d,d}vecA$. By property P7 this equation is satisfied if and only if A is a symmetric matrix. ■

Some caution is needed when dealing with third cumulants which are rank deficient, that is when some singular values equal zero. Left singular vectors corresponding to null singular values might or might not be represented as vectorized, real and symmetric matrices. The following example illustrates the point. Let $K_3 = \delta \otimes \delta^T \otimes \delta$, where $\delta \in \mathbb{R}_0^d$, be the third cumulant of a $d-$dimensional random vector x. This might happen, for example, when x is either a skew-normal random vector (see, for example, De Luca and Loperfido, 2012) or a location mixture of two normal distributions (Loperfido, 2013). Also, let v_1, ..., v_{d-1} be $d-$dimensional real vectors satisfying $v_i^T\delta = 0$, $v_i^T v_j = 0$ and $v_i^T v_j = 1$ for $i \neq j$ and $i, j = 1, ..., d-1$. It easily follows that left singular vectors corresponding to null singular values might be either of the form $vec\left(\lambda v_i^T\right) = \lambda \otimes v_i$, $vec\left(v_i^T\lambda\right) = v_i \otimes \lambda$, $vec\left(v_j v_i^T\right) = v_j \otimes v_i$ or $vec\left(v_i v_i^T\right) = v_i \otimes v_i$, with $\lambda = \delta / \left\|\delta\right\|$, $i \neq j$ and $i, j = 1, ..., d-1$. Only vectors of the latter form are vectorized, symmetric real matrices.

The next theorem builds upon the previous one to obtain another decomposition for the third multivariate moment.

Theorem 2 *The third moment of a $d-$dimensional random vector might be represented as $\sigma_1 V_1 \otimes u_1 + ... + \sigma_d V_d \otimes u_d$, where $\left\|V_i\right\| = \left\|u_i\right\| = 1$, $tr\left(V_i V_j\right) = u_i^T u_j = 0$, $V_i = V_i^T \in \mathbb{R}^d \times \mathbb{R}^d$, $u_i \in \mathbb{R}^d$, $\sigma_i \in \mathbb{R}_+$, for $i \neq j$ and $i, j = 1, ..., d$.*

Proof. Let M_3 be the third moment of the $d-$dimensional random vector $x = \left(X_1, ..., X_d\right)^T$. We have $M_3 = E\left(x \otimes x^T \otimes x\right)$ by definition, that is $M_3 = E\left(x \otimes x^T \otimes x\right)$ by property P3. Hence we can represent M_3 as the block column vector $\left(B_1, ..., B_d\right)^T$, where $B_i = E\left(X_i xx^T\right) \in \mathbb{R}^d \times \mathbb{R}^d$, $i = 1, ..., d$. The matrix M_3 might also be represented as $E\left(x \otimes x \otimes x^T\right)$ by recalling property P3, so that $M_3 = \left[E\left(X_1 x \otimes x\right), ..., E\left(X_d x \otimes x\right)\right]$. The vector $E\left(X_i x \otimes x\right)$ is just the vectorized matrix B_i, since $vec\left(xx^T\right) = x \otimes x$. We conclude that the $i-$th column of M_3 is the vectorized $d \times d$ matrix B_i. Equivalently, M_3^T is the rearrangement of M_3, as defined by Van Loan and Pitsianis (1993). They also showed that any block matrix A admits the representation $B_1 \otimes C_1 + ... + B_r \otimes C_r$ if and only if its permuted version admits the representation $vec(B_1) \otimes vec^T(C_1) + ... + vec(B_r) \otimes vec^T(C_r)$. By Theorem 1, the singular value decomposition of the third moment is $\sigma_1 vec(V_1) u_1^T + ... + \sigma_r vec(V_r) u_r^T$, where V_i is a $d \times d$ symmetric matrix. Apply now ordinary properties of transposition and vectorization to obtain $M_3^T = \sigma_1 vec(u_1) \otimes vec^T(V_1) + ... + \sigma_r vec(u_r) \otimes vec^T(V_r)$. By the above mentioned propertiy of rearrangements we have $M_3 = \sigma_1 u_1 \otimes V_1 + ... + \sigma_r u_r \otimes V_r$. By definition, $vec(V_i)$ and u_i are the left and right singular vectors of M_3, satisfying $vec^T(V_i) vec(V_i) = u_i^T u_i = 1$ and $vec^T(V_i) vec(V_j) = u_i^T u_j = 0$ for $i, j = 1, ..., n$ and $i \neq j$. Apply now property P4 to show that $\|V_i\| = \|u_i\| = 1$ and $tr(V_i V_j) = u_i^T u_j = 0$ for $i, j = 1, ..., n$ and $i \neq j$. ■

3 Skewness

This section uses the singular value decomposition of the third standardized moment to investigate some relationships between well-known measures of multivariate skewness. The third standardized moment (or cumulant) of the random vector x is the the third moment of $z = (Z_1, ..., Z_d)^T = \Sigma^{-1/2}(x - \mu)$, where $\Sigma^{-1/2}$ is the inverse of the positive definite square root $\Sigma^{1/2}$ of $\Sigma = var(x)$, which is assumed to be positive definite: $\Sigma^{1/2} = \left(\Sigma^{1/2}\right)^T$, $\Sigma^{1/2} > 0$ and $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$.

Mardia (1970) defined the skewness of a random vector x as $\beta_{1,d}^M(x) = E\left[\left(z^T w\right)^3\right]$, where $w = \Sigma^{-1/2}(y - \mu)$, z is the same as above while x and y are two $d-$dimensional, independent and identically distributed random vectors. Mardia's skewness is by far the most popular measure of multivariate skewness. Its statistical applications include multivariate normality testing (Mardia, 1970), assessing the robustness of MANOVA statistics (Davis, 1980) and density approximation (Van Hulle, 2005).

Another measure of multivariate skewness is $\beta_{1,d}^P(x) = E\left(z^T z z^T w w^T w\right)$, where z and w are the same as above. It has been independently proposed by several authors (Davis, 1980; Isogai, 1983; Mori *et al*, 1993). In this paper, we shall refer to $\beta_{1,d}^P(x)$ as to partial skewness, to remind that it does not depend on the cross-product moment $E(Z_i Z_j Z_k)$ when i, j, k differ from each other. Partial skewness is far less popular than Mardia's skewness. Like the latter measure, however, it has been applied to multivariate normality testing (Henze 1997) and to the assessment of the robustness of MANOVA statistics (Davis, 1980).

Malkovich and Afifi (1973) defined the multivariate skewness of a random vector x as the maximum value $\beta_{1,d}^D(x)$ attainable by $\beta_1\left(c^T x\right)$, where c is a nonnull, $d-$dimensional real vector and $\beta_1(Y)$ is the squared third standardized moment of the random variable Y. In this paper, we shall refer to $\beta_{1,d}^D(x)$ as to directional skewness, to remind that it is the maximum attainable skewness by a projection of the random vector x onto a direction. Statistical applications of directional skewness include normality testing (Malkovich and Afifi, 1973), projection pursuit and cluster analysis (Loperfido, 2013).

The above mentioned measures of multivariate skewness have been mainly used for testing multivariate normality, are invariant with respect to one-to-one affine transformations and generalize to the multivariate case a widely used measure of univariate skewness. To the best of our knowledge, no one investigated the relationships between Mardia's, partial and directional skewness. The following theorem addresses the problem by means of inequalities.

Theorem 3 *Let $\beta_{1,d}^{M}$, $\beta_{1,d}^{P}$ and $\beta_{1,d}^{D}$ be the Mardia's skewness, partial skewness and directional skewness of a $d-$dimensional random vector x. Then they satisfy the inequalities $\beta_{1,d}^{D} \leq \beta_{1,d}^{M}$ and $\beta_{1,d}^{P}/d \leq \beta_{1,d}^{M}$.*

Proof. Without loss of generality we can assume that x is a standardized random vector, so that its third moment M_3 is also its third standardized cumulant. Also, let v_i and u_i be the left and right singular vectors associated with the $i-$th singular value σ_i, with $i = 1, ..., d$ and $\sigma_1 \geq ... \geq \sigma_d$. From Theorem 1 we know that $v_i = vec(V_i)$, where $V_i = V_i^T \in \mathbb{R}^d \times \mathbb{R}^d$, $i = 1$, ..., d. Mardia's skewness of x might be represented as the squared Frobenius norm of the third standardized cumulant: $\beta_{1,d}^{M}(x) = \|M_3\|^2 = tr\left(M_3 M_3^T\right)$. Property P4 implies that $\beta_{1,d}^{M} = \sigma_1^2 + ... + \sigma_d^2$. Directional skewness is the maximum value attained by the square of $(c \otimes c)^T M_3 c$ over the set of $d-$dimensional vectors of unit length (Loperfido, 2013). The bilinear form $a^T M_3 b$, where $a \in \mathbb{R}^{d^2}$, $b \in \mathbb{R}^d$ and $\|a\| = \|b\| = 1$, attains its maximum value σ_1 when $a = v_1$ and $b = u_1$, by elementary properties of the singular value decomposition. It easily follows that $\beta_{1,d}^{D} \leq \sigma_1^2 \leq \beta_{1,d}^{M}$ and this completes the first part of the proof.

Mori *et al* (1993) introduced the vector-valued measure of skewness $\alpha = E\left(x^T x x\right)$, whose $i-$th component is the sum $E\left(X_i^2 X_1\right) + ... + E\left(X_i^2 X_d\right)$, for $i = 1$, ..., d. Equivalently, $\alpha_i = tr(B_i)$, where $B_i = E\left(X_i x x^T\right)$. Basic properties of vectorized matrices imply that $\alpha_i = tr(B_i) = vec^T(I_d)\, vec(B_i)$. From the proof of Theorem 2 we know that M_3 might be represented as $[vec(B_1), ..., vec(B_d)]$. Hence we have $\alpha = M_3^T vec(I_d)$. The squared norm of α equals $\beta_{1,d}^{P}$: $\beta_{1,d}^{P} = vec^T(I_d)\, M_3 M_3^T vec(I_d)$ (Mori *et al*, 1993). By Theorem 2 we have $M_3^T vec(I_d) = \sigma_1\left(V_1^T \otimes u_1^T\right) vec(I_d) + ... + \sigma_d\left(V_d^T \otimes u_d^T\right) vec(I_d)$, which can be simplified into $\sigma_1 V_1 u_1 + ... + \sigma_d V_d u_d$ by recalling that $vec(a) = a$ for any vector a and property P5. Then $\left\|M_3^T vec(I_d)\right\|^2$ of $M_3^T vec(I_d)$ is $vec^T(I_d)\, M_3 M_3^T vec(I_d)$, that is the sum of all products $\sigma_i \sigma_j u_j^T V_j V_i u_i$, for $i, j = 1$, ..., d. By definition, we have $\|V_i\| = \|u_i\| = 1$. Hence the norm of the vector $V_i u_i$ is never greater than 1: $\|V_i u_i\| = u_j^T V_j V_i u_i \leq 1$. As a direct consequence, $\left\|M_3^T vec(I_d)\right\|^2$ is never greater than the sum of the products $\sigma_i \sigma_j$, for $i, j = 1$, ..., d, that is $(\sigma_1 + ... + \sigma_d)^2$. A squared sum of real values is never greater than the sum of the squared values themselves, multiplied by their number $(\sigma_1 + ... + \sigma_d)^2 \leq d\left(\sigma_1^2 + ... + \sigma_d^2\right)$. We shall now complete the proof by recalling that Mardia's skewness equals the trace of $M_3 M_3^T$: $\left\|M_3^T vec(I_d)\right\|^2 \leq d \cdot \beta_{1,d}^{M}$ and $\beta_{1,d}^{P}/d \leq \beta_{1,d}^{M}$. ■

A natural question to ask is whether Mardia's skewness, partial skewness and directional skewness might be equal to one another. The answer is in the affirmative, as shown in the following theorem.

Theorem 4 *Mardia's skewness, directional skewness and partial skewness of a random vector whose third cumulant is a rank one matrix are equal to each other.*

Proof. Without loss of generality we can assume that $x = (X_1, ..., X_d)^T$ is a standardized random vector, so that its third moment M_3 is also its third standardized cumulant. As shown in the proof of Theorem 2, M_3 is the block column vector $(B_1, ..., B_d)^T$, where $B_i = E\left(X_i x x^T\right) \in \mathbb{R}^d \times \mathbb{R}^d$, $i = 1, ..., d$. By assumption, M_3 is a matrix of rank one, meaning that it has only a nonzero singular value and admits the factorization $\sigma v u^T$, with $\sigma \in \mathbb{R}$, $u \in \mathbb{R}^d$, $v \in \mathbb{R}^{d^2}$ and $\|v\| = \|u\| = 1$. Without loss of generality, v might be represented as a vectorized, symmetric, $d \times d$ matrix A whose columns are the vectors a_i: $v = vec(A)$, $A = (a_1, ..., a_d)$, $a_i \in \mathbb{R}^d$. As a direct consequence, we have $B_i = a_i u^T$, which might be simplified into $B_i = c_i u\, u^T$, with $c_i \in \mathbb{R}$, by noticing that B_i is a symmetric matrix. Since the columns of A are proportional to each other, it admits the factorization uc^T, where $c = (c_1, ..., c_d)^T$. By Theorem 1, A is a symmetric matrix, so that $c = u$, $v = vec\left(u\, u^T\right) = u \otimes u$ and $M_3 = u \otimes u \otimes u^T$.

We shall now consider the three measures of multivariate skewness, beginning with Mardia's skewness. From the proof of Theorem 3 we know that it equals the sum of the squared singular values of M_3. By assumption, M_3 is a matrix of rank one so that Mardia's skewness equals the square of the only nonzero singular value: $\beta_{1,d}^M = \sigma^2$. We shall now consider partial skewness, that is $\beta_{1,d}^P(x) = vec^T(I_d)\, M_3 M_3^T vec(I_d)$. Property P5 leads us to $M_3^T vec(I_d) = \sigma\left(u^T u\right) u$. As a direct consequence, partial skewness is $\beta_{1,d}^P(x) = \sigma^2$, that is Mardia's skewness, thus completing the second part of the proof. We shall now consider the directional skewness of x, that is the maximum of $\beta_1\left(c^T x\right)$ over the set of all $d-$dimensional vectors of unit length, where $\beta_1(Y)$ is the squared third standardized moment of the random variable Y. By the above mentioned properties of the Kronecker product and the definition of M_3 we have $(a \otimes a)^T M_3 a = \sigma\left(u^T a\right)^3$, where $a \in \mathbb{R}^d$, which attains its maximum when a and u are proportional to each other. From the proof of Theorem 3 we know that directional skewness is the maximum value attained by the square of $(c \otimes c)^T M_3 c$ over the set of $d-$dimensional vectors of unit length, so that $\beta_{1,d}^D = \sigma^2$, which again equals Mardia's skewness. ∎

Random vectors whose third standardized moments have only one nonzero singular value appear in hidden truncation models, finite mixture models and multivariate density approximations (Christiansen and Loperfido, 2014).

4 Independence

Results in this section deal with the third moment's SVD under independence assumptions. Theorems 5 and 6 deal with left singular vectors and singular values, respectively. Some possible applications to inference for linear processes are briefly sketched at the end of the section.

Theorem 5 *Let $x = (X_1, ..., X_d)^T$ be a vector of mutually independent random variables with finite third moments. Also, let $y = Ax$ be a linear transformation of x, where A is a $d \times d$, nonsingular matrix. Then left singular vectors of the third standardized cumulant of y are vectorized, symmetric matrices of rank one.*

Proof. Without loss of generality we can assume that X_1, ..., X_d are standardized random variables: $E(X_i) = 0$ and $var(X_i) = 1$ for $i = 1$, ..., d. Let $D = diag\left(\gamma_1^2, ..., \gamma_d^2\right)$, where $E\left(X_i^3\right) = \gamma_i$ for $i = 1$, ..., d, be the diagonal matrix whose $i-$th diagonal element is the squared third moment of X_i. Also, let E_{ij} be the $d \times d$ matrix whose only nonzero element is one and belongs to the $i-$th row and the $j-$th column of E_{ij}, for $i, j = 1$, ..., d. By assumption $X_1, ..., X_d$ are mutually independent, so that $E(X_i X_j X_k) = 0$ whenever either $i \neq j$, $i \neq k$ or $k \neq j$. As a direct consequence, the third moment of x might be represented as a block column vector: $M_{3,x} = \{\gamma_i E_{ii}\}$, for $i = 1$, ..., d. Let $z = \Sigma^{-1/2} y$, where $\Sigma^{-1/2}$ is the inverse of the positive definite square root $\Sigma^{1/2}$ of $var(y) = \Sigma = AA^T$: $\Sigma^{1/2} = \left(\Sigma^{1/2}\right)^T$, $\Sigma^{1/2} > 0$ and $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$. The standardized random vectors z and x are one-to-one linear functions of each other, so that there exists a $d \times d$ orthogonal matrix H such that $z = Hx$. The third moment of z is then $M_{3,z} = \left(H^T \otimes H^T\right) M_{3,x} H$ (De Luca and Loperfido, 2012). It follows that $M_{3,z}^T M_{3,z} = H^T M_{3,x}^T (H \otimes H) \left(H^T \otimes H^T\right) M_{3,x} H = H^T D H$, by the definition of D, well-known properties of the Kronecker product and orthogonality of H. Then the $i-$th singular value of $M_{3,z}$ and the associated right singular vector are $|\gamma_i|$ and h_i, where h_i is the $i-$th column of H. Similarly, we have $M_{3,z} M_{3,z}^T = \left(H^T \otimes H^T\right) diag\left[vec(D)\right] (H \otimes H)$, where $diag\left[vec(D)\right]$ is the $d^2 \times d^2$ diagonal matrix whose diagonal elements are those in the vectorized matrix D. The matrix $M_{3,z} M_{3,z}^T$ can then be decomposed into the sum of all products $d_{ij} (h_i \otimes h_i) \left(h_j^T \otimes h_j^T\right)$, for $i, j = 1$, ..., d, where d_{ij} is the element on the $i-$th row and the $j-$th column of the matrix D. By definition, $d_{ii} = \gamma_i^2$ for $i = 1$, ..., d and $d_{ij} = 0$ for $i \neq j$, $i, j = 1$, ..., d, thus leading to the simpler factorization $M_{3,z} M_{3,z}^T = \gamma_1^2 (h_1 \otimes h_1) \left(h_1^T \otimes h_1^T\right) + ... + \gamma_d^2 (h_d \otimes h_d) \left(h_d^T \otimes h_d^T\right)$. We conclude that the left singular vector of $M_{3,z}$ associated with the singular value $|\gamma_i|$ is $h_i \otimes h_i$, which is a vectorized, symmetric matrix of rank one. ■

....

Theorem 6 *Let $y = Ax$, where A is a $d \times d$, nonsingular matrix and $x = (X_1, ..., X_d)^T$ is a vector of mutually independent random variables. Also, let γ_i be the third standardized moment of X_i, for $i = 1$, ..., d. Then the directional skewness of y is $\beta_{1,d}^D = \max\left(\gamma_1^2, ..., \gamma_d^2\right)$, while Mardia's skewness and partial skewness of y are $\beta_{1,d}^M = \beta_{1,d}^P = \gamma_1^2 + ... + \gamma_d^2$.*

Proof. The notation in this proof is the same as in the previous one. We shall first consider Mardia's skewness. The identity $M_{3,z}M_{3,z}^T = \gamma_1^2 h_1 h_1^T \otimes h_1 h_1^T + ... + \gamma_d^2 h_d h_d^T \otimes h_d h_d^T$ follows from $M_{3,z}M_{3,z}^T = \gamma_1^2 (h_1 \otimes h_1)(h_1^T \otimes h_1^T) + ... + \gamma_d^2 (h_d \otimes h_d)(h_d^T \otimes h_d^T)$ in the previous proof and property P2. Recall now property P6 and that Mardia's skewness equals the trace of $M_3 M_3^T$, implying that $\beta_{1,d}^M = \gamma_1^2 tr^2 (h_1 h_1^T) + ... + \gamma_d^2 tr^2 (h_d h_d^T) = \gamma_1^2 (h_1^T h_1)^2 + ... + \gamma_d^2 (h_d^T h_d)^2$. Since h_1, ... h_d are real vectors of unit norm Mardia's skewness is $\beta_{1,d}^M = \gamma_1^2 + ... + \gamma_d^2$ and this completes the first part of the proof.

We shall now consider partial skewness. We have $(h_i^T \otimes h_i^T) vec(I_d) = h_i^T h_i$ for $i = 1, ..., d$ by property P5. It follows that $\beta_{1,d}^P = vec^T (I_d) M_3 M_3^T vec(I_d) = \gamma_1^2 (h_1^T h_1)^2 + ... + \gamma_d^2 (h_d^T h_d)^2 = \gamma_1^2 + ... + \gamma_d^2$ and this completes the second part of the proof.

We shall now consider directional skewness $\beta_{1,d}^D$, that is $\left[(c \otimes c)^T M_{3,z} c\right]^2$ for an appropriate $d-$dimensional vector c of unit length (Loperfido, 2013). The singular value decomposition $M_{3,z} = |\gamma_1| (h_1 \otimes h_1) h^T + ... + |\gamma_d| (h_d \otimes h_d) h^T$ implies that $(c \otimes c)^T M_{3,z} c = |\gamma_1| (c^T h_1)^3 + ... + |\gamma_d| (c^T h_d)^3$. Without loss of generality we may assume that c maximizes $(c \otimes c)^T M_{3,z} c$. Without loss of generality we can also assume that $\gamma_1 > 0$ and $|\gamma_1| \geq |\gamma_i|$ for $i = 2, ..., d$. By definition, $h_1, ..., h_d$ are unit-length vectors which constitute an orthogonal basis for $\mathbb{R}^d$, so that $c = w_1 h_1 + ... + w_d h_d$ for some w_1, ..., w_d satisfying $w_1^2 + ... + w_d^2 = 1$. It also implies that $(c \otimes c)^T M_{3,z} c = |\gamma_1| w_1^3 + ... + |\gamma_d| w_d^3$, which is maximized if and only if $w_1 = 1$ and $w_i = 0$ for $i = 2, ..., d$. As a direct consequence, $\beta_{1,d}^D = \gamma_1^2$. ■

We shall now hint to a possible application of the above results to inference for linear processes. Suppose we want to check whether the random process $\{Y_t, t \in \mathbb{Z}\}$ might be represented by means of the linear transformations $y_n = (Y_n, ..., Y_{n+k-1})^T = A_n x_n$ for some sequence $\{A_n, n \in \mathbb{N}\}$ of $k \times k$, nonsingular matrices A_n and for a random vector $x_n = (X_n, ..., X_{n+k-1})^T$ belonging to a random process $\{X_t, t \in \mathbb{Z}\}$, whose components are independent random variables satisfying $E(|X_t|) < +\infty$. Such a process has been investigated by Chen and Deo (2004) within the context of time series analysis.

To check whether $\{W_t, t \in \mathbb{Z}\}$ admits the above representation we might compare Mardia's skewness and partial skewness of its realizations: if they are significantly different we reject the corresponding hypothesis. Similarly, we might use left eigenvectors of third cumulants of standardized realizations. Sampling distributions might be obtained by means of bootstrap methods for dependent data. Technical developments of the proposed inferential procedure are definitely outside the scope of the present paper, which is focused on algebraic statistics.

5 Invariance

This section investigates the singular value decomposition of the third moment when the underlying distribution satisfies the invariance assumptions of either exchangeability or reversibility.

A random vector is said to be exchangeable if its distribution is invariant with respect to permutations of its elements' indices. Loperfido *et al* (2007) and Loperfido and Guttorp (2008) discuss applications of skewed, exchangeable random vectors to reliability theory and environmental sciences, respectively.

When moments of appropriate order exist, mean and variance of an exchangeable random vector are proportional to a vector of ones and a correlation matrix with identical off-diagonal elements. The structure of its third moment is slightly more complicated. In the general case, the third moment of a $d-$dimensional random vector contains at most $d(d+1)(d+2)/6$ distinct elements. However, when the vector itself is assumed to be exchangeable, the number of distinct elements never exceeds 3.

A matrix is totally symmetric if its diagonal (extra-diagonal) elements are all equal. At least one eigenvector of the fourth moment matrix of an exchangeable random vector is a vectorized, totally symmetric matrix (Loperfido, 2011). The following theorem shows that a similar result also holds for the third moment.

Theorem 7 *Let M_3 be the third moment of a $d-$dimensional, exchangeable random vector. Then there is at most one singular value of M_3 which is different from all other singular values, its right singular vector is proportional to a vector of ones and its left singular vector is a vectorized, totally symmetric matrix.*

Proof. Let P_π be the $d \times d$ permutation matrix corresponding to a permutation π of the first d positive integers. Hence the components of $P_\pi x$ are the components of x, rearranged according to π, and the third moment of $P_\pi x$ is $M_{3,\pi} = (P_\pi \otimes P_\pi) M_3(x) P_\pi^T$ (see, for example, De Luca and Loperfido, 2012). The matrix $M_{3,\pi}^T M_{3,\pi} = P_\pi M_3^T (P_\pi \otimes P_\pi)^T (P_\pi \otimes P_\pi) M_3 P_\pi^T$ might be simplified into $M_{3,\pi}^T M_{3,\pi} = P_\pi M_3^T M_3 P_\pi^T$, since the inverse of $(P_\pi \otimes P_\pi)$ is $\left(P_\pi^T \otimes P_\pi^T\right) = (P_\pi \otimes P_\pi)^T$, by orthogonality of P_π and property P2. The exchangeability assumption implies that $P_\pi x$ and x are identically distributed, with identical third moments: $M_3 = M_{3,\pi}$. As a direct consequence, we have $M_3^T M_3 = P_\pi M_3^T M_3 P_\pi^T$ for any $d \times d$ permutation matrix P_π, which is possible if and only if $M_3^T M_3$ is a $d \times d$ totally symmetric matrix. As such, one of its eigenvectors is the $d-$dimensional vector of ones $1_d = (1, ..., 1)^T \in \mathbb{R}^d$. The associated eigenvalue is always different from other eigenvalues of M_3, but when $M_3^T M_3$ is a null matrix (that is when M_3 is a null matrix, too) and when $M_3^T M_3$ is a diagonal matrix. We have then proved the first two parts of the theorem.

We shall now prove the third part of the theorem. By definition, the left singular vector y of M_3 associated with the same singular value σ of the right singular vector 1_d satisfies $M_3 1_d = \sigma y$. The above mentioned exchangeability argument then implies $(P_\pi \otimes P_\pi) M_3 P_\pi^T 1_d = \sigma y$, which simplifies to $(P_\pi \otimes P_\pi) M_3 1_d = \sigma y$, since 1_d is a vector of ones. Apply now ordinary properties of the permutation matrix and the Kronecker product to obtain $(P_\pi \otimes P_\pi)^{-1} = P_\pi^{-1} \otimes P_\pi^{-1} = P_\pi^T \otimes P_\pi^T$ and hence $M_3 1_d = \sigma \left(P_\pi^T \otimes P_\pi^T\right) y$. By Theorem 1 in Section 2, the d^2-dimensional vector y might be represented as a vectorized, symmetric matrix Y: $y = vec(Y) = vec\left(Y^T\right)$. Also, recall property P5 and write $M_3 1_d = \lambda vec\left(P_\pi Y P_\pi^T\right)$ for any $d \times d$ permutation matrix P_π. Hence we have $Y = P_\pi Y P_\pi^T$ for any $d \times d$ permutation matrix P_π, which holds true if and only if Y is a $d \times d$ totally symmetric matrix. ■

Another invariance property holds true for a reversible process $\{X_t,\ t \in \mathbb{Z}\}$, since the random vectors $(X_{t_1}, \ldots, X_{t_n})^T$ and $(X_{\tau - t_1}, \ldots, X_{\tau - t_n})^T$ have the same probability distribution, for all τ, $t_1, \ldots, t_n \in \mathbb{Z}$. With a little abuse of language and for the sake of simplicity, we shall refer to a vector $(X_t, ..., X_{t+n})^T$ of random variables from a reversible random process as to a reversible random vector. A reversible process is also stationary, meaning that vectors $(X_t, ..., X_{t+n})^T$ and $(X_{t+h}, ..., X_{t+n+h})^T$ are identically distributed, for $t, h \in \mathbb{Z}$ and $n \in \mathbb{N}$. Stationarity is particularly useful when dealing with moments. For example, it implies that $E(X_t) = E(X_{t+h})$ for $t, h \in \mathbb{Z}$, so that the expected value of $(X_t, ..., X_{t+n})$ is proportional to a $(n+1)-$dimensional vector of ones. Similarly, stationarity implies that $E(X_t X_{t+h})$ only depends on h, thus making the covariance matrix of $(X_t, ..., X_{t+n})^T$ a symmetric Toeplitz matrix. Interesting enough, the first, second and third moments of a reversible random vector are completely identified by a scalar, a vector and a matrix, respectively.

The following theorem relates left singular vectors of the third moment to bisymmetric matrices, that is square matrices which are symmetric about both of their main diagonals.

Theorem 8 *Let M_3 be the third moment of a $d-$dimensional random vector belonging to a reversible process. Also, let the positive singular values of M_3 be different from each other. Then left singular vectors of M_3 are vectorized, bisymmetric matrices.*

Proof. Let $x = (X_1, ..., X_d)^T$ be a $d-$dimensional, reversible random vector whose third moment is M_3. Also, let J_d be the $d \times d$ exchange matrix, that is the $d \times d$ permutation matrix whose nonzero entries reside on the counterdiagonal (Rao and Rao, 1998, page 502). The product $J_d x$ contains the same elements of x, arranged in reverse order: $J_d x = (X_d, ..., X_1)^T$. The third moment of $J_d x$ is $(J_d \otimes J_d) M_3 J_d^T$, by linear properties of the third moment (see, for example, De Luca and Loperfido, 2012). The random vectors that x and $J_d x$ are identically distributed, due to the reversibility of x. It follows that x and $J_d x$ have identical third moments: $M_3 = (J_d \otimes J_d) M_3 J_d^T$.

Let σ_i be a positive singular value corresponding to the left singular vector l_i of M_3: $M_3 M_3^T l_i = \sigma_i l_i$, for $i = 1, ..., d$. Since $M_3 = (J_d \otimes J_d) M_3 J_d^T$, we can also write $(J_d \otimes J_d) M_3 J_d^T J_d M_3^T \left(J_d^T \otimes J_d^T\right) l_i = \sigma_i l_i$. Equivalently, by orthogonality of J_d and by properties P1 and P2, we have $M_3 M_3^T \left(J_d^T \otimes J_d^T\right) l_i = \sigma_i \left(J_d^T \otimes J_d^T\right) l_i$. It follows that the d^2-dimensional vector l_i is the left singular vector of M_3 corresponding to the positive singular value σ_i if and only the

d^2-dimensional vector $\left(J_d^T \otimes J_d^T\right) l_i$ is a left singular vector of M_3 corresponding to the same singular value. This is possible only if $l_i = \left(J_d^T \otimes J_d^T\right) l_i$, by the assumption that all positive singular values differs from each other. The identity $l_i = \left(J_d^T \otimes J_d^T\right) l_i$ is equivalent to the following one: $B_i = J_d B_i J_d^T$, by letting $l_i = vec\left(B_i\right)$, where B_i is a $d \times d$ matrix, and by applying property P3. By Theorem 1, B_i is also a symmetric matrix, so that $B_i = J_d B_i^T J_d$. Since this identity characterizes persymmetric matrices (Rao and Rao, 1998, page 503) B_i is at the same time symmetric and persymmetric, that is bisymmetric. ■

6 Fourth moment

A natural question to ask is whether results in the previous sections might be extended to higher-order moments. A complete answer is beyond the scope of this paper, but the following theorems suggest that it might be in the affirmative, at least for fourth-order moments. The theorems and their proofs are very similar to those in Section 2, thus hinting the existence of a general procedure for dealing with higher-order moments.

Let $x = (X_1, ..., X_d)^T$ be a $d-$dimensional random vector satisfying $E\left(X_i^4\right) < +\infty$, for $i = 1, ..., d$. The fourth moment of x is the $d^2 \times d^2$ matrix $M_4 = E\left(x \otimes x^T \otimes x \otimes x^T\right)$. The following theorem shows that any fourth moment can be represented as the sum of tensor products of symmetric, real matrices.

Theorem 9 *The fourth moment M_4 of a $d-$dimensional random vector admits the representation $\lambda_1 \Omega_1 \otimes \Omega_1 + ... + \lambda_r \Omega_r \otimes \Omega_r$, where $r = rank\left(M_4\right)$, $\|\Omega_i\| = 1$, $tr\left(\Omega_i \Omega_j\right) = 0$, $\Omega_i = \Omega_i^T \in \mathbb{R}^d \times \mathbb{R}^d$, $\lambda_i \in \mathbb{R}_+$, for $i \neq j$ and $i, j = 1, ..., r$.*

Proof. Let $x = (X_1, ..., X_d)^T$ be a random vector whose fourth moment is $M_4 = \left\{\mu_{ijhk}\right\}$, where $\mu_{ijhk} = E\left(X_i X_j X_h X_k\right)$ for $i, j, h, k = 1, ..., d$. Since M_4 is $E\left(x \otimes x^T \otimes x \otimes x^T\right)$ by definition, it also equals $E\left(x \otimes x \otimes x^T \otimes x^T\right)$, by property P3. It follows that M_4 is a matrix composed by d^2 d^2-dimensional row vectors v_k^T, stacked upon each other:

$$M_4 = \begin{pmatrix} v_1^T \\ ... \\ v_{d^2}^T \end{pmatrix} \tag{a}$$

The $k-$th vector v_k has the form $E\left(X_i X_j x \otimes x\right)$, where $i = 1 + [k/d]$, $j = 1 + k - [k/d]$ and $i = [k/d]$ is the integer part of k/d, for $k = 1, ..., d^2$. The fourth moment M_4 might also be represented as a block matrix $\{M_{pq}\}$, where $M_{pq} = \left\{E\left(X_p X_q x x^T\right)\right\}$ is a $d \times d$ symmetric matrix. The vector $vec\left(M_{ij}\right)$ is

$$vec\left\{E\left(X_i X_j x x^T\right)\right\} = E\left\{X_i X_j vec\left(x x^T\right)\right\} = E\left\{X_i X_j\left(x \otimes x\right)\right\} = v_k,$$

implying that M_4 coincides with its rearrangement, as defined by Van Loan and Pitsianis (1993). The fourth moment is a symmetric, positive semidefinite matrix (see, for example, Loperfido, 2011), so that it admits the representation

$M_4 = \lambda_1 u_1 u_1^T + ... + \lambda_r u_r u_r^T$, where λ_1, ..., λ_r are the positive eigenvalues of M_4 and u_1, ..., u_r are the corresponding eigenvectors, which are of unit norm and mutually orthogonal. Van Loan and Pitsianis (1993) showed that any block matrix A admits the representation $B_1 \otimes C_1 + ... + B_r \otimes C_r$ if and only if its permuted version admits the representation $vec(B_1) \otimes vec^T(C_1) + ... + vec(B_r) \otimes vec^T(C_r)$. Since M_4 coincides with its rearrangement and u_1, ..., u_r are vectorized, symmetric matrices (Loperfido, 2011) we may also write $M_4 = \lambda_1 vec(\Omega_1) \otimes vec^T(\Omega_1) + ... + \lambda_r vec(\Omega_r) \otimes vec^T(\Omega_r)$, where $vec(\Omega_i) = u_i$, $\Omega_i = \Omega_i^T \in \mathbb{R}^d \times \mathbb{R}^d$ for $i = 1, ..., r$ and $\|\Omega_i\| = 1$. Property P4 implies that $\|\Omega_i\| = 1$ and $tr(\Omega_i \Omega_j) = 0$ for $i, j = 1, ..., n$ and $i \neq j$. ■

The fourth moment of a random vector is closely related to the fourth cumulant of the random vector itself. More precisely, let $x = (X_1, ..., X_d)^T$ be a $d-$dimensional random vector satisfying $E\left(X_i^4\right) < +\infty$, for $i = 1, ..., d$. Its fourth cumulant is the $d^2 \times d^2$ block matrix $K_4 = \{M_{pq}\}$, where $M_{pq} = \log E\left[\exp\left(\iota t^T x\right)\right] / \partial t_p \partial t_q \partial t \partial t^T$, $\iota = \sqrt{-1}$ and $t^T = (t_1, ..., t_d)$, for $p, q = 1, ..., d$. Equivalently, the element in the $i-$th row and in the $j-$th column of the $(h, k)-$th block is $\kappa_{ijhk} = \log E\left[\exp\left(\iota t^T x\right)\right] / \partial t_i \partial t_j \partial t_h \partial t_k$, for $i, j, h, k = 1, ..., d$. Loperfido (2011) showed that eigenvectors associated to positive eigenvalues of the fourth moment of a $d-$dimensional random vectors are vectorized, symmetric real matrices. The following theorem shows that a similar property holds for fourth cumulants.

Theorem 10 *Eigenvectors corresponding to nonzero eigenvalues of fourth multivariate cumulants are vectorized, symmetric matrices.*

Proof. The fourth cumulant of a $d-$dimensional random vector x with mean μ and variance Σ is $K_4 = \overline{M}_4 - (I_{d^2} + K_{d,d})(\Sigma \otimes \Sigma) - vec(\Sigma) vec^T(\Sigma)$, where $\overline{M}_4$ is the fourth centered moment of x (see, for example, Loperfido, 2011). First, recall the identity $\overline{M}_4 = K_{d,d} \overline{M}_4$, from the first part of Theorem 2 in Loperfido (2011). Second, use property P7 and symmetry of the covariance matrix to show that $vec(\Sigma) = K_{d,d} vec(\Sigma)$. The identity $K_{d,d} K_{d,d} = I_{d^2}$ follows from properties P8 and P9. We can then write $K_4 = K_{d,d} K_4$ and $K_{d,d} K_4 v = \lambda v$, where λ is the nonzero eigenvalue of K_4 corresponding to the eigenvector v: $K_4 v = \lambda v$. Since v is a d^2- dimensional vector, it may be represented as a vectorized, $d \times d$ matrix A. The matrix $K_{d,d}$ is at the same time symmetric and orthogonal, so that $K_4 v = \lambda K_{d,d} v$. The above equations lead to the following one: $v = K_{d,d} v$, which can be represented as $vec(A) = K_{d,d} vec(A)$. By property P7 this equation is satisfied if and only if A is a symmetric matrix. ■

Corollary 11 *The fourth cumulant K_4 of a $d-$dimensional random vector admits the representation $\lambda_1 \Omega_1 \otimes \Omega_1 + ... + \lambda_r \Omega_r \otimes \Omega_r$, where $r = rank(K_4)$, $\|\Omega_i\| = 1$, $tr(\Omega_i \Omega_j) = 0$, $\Omega_i = \Omega_i^T \in \mathbb{R}^d \times \mathbb{R}^d$, $\lambda_i \in \mathbb{R}$, for $i \neq j$ and $i, j = 1, ..., r$.*

The proof is very similar to the first one in this section and is therefore omitted.

7 Numerical example

According to the volatility feedback theory (French *et al*, 1987) an effect of news on stock prices (direct effect) determines a negative effect (feedback effect) on the prices themselves by making them more volatile. The multivariate SGARCH model (De Luca *et al*, 2006) quantifies the feedback effects in several stock markets. A major implication of the model is that third cumulants of market price innovations in follower markets, conditionally on either bad or good news from leader markets, are matrices of rank one. Third cumulants of unconditional market innovations are matrices of rank one only when the ratio of the feedback effect to the direct effect remains constant across markets (De Luca and Loperfido, 2012).

We shall use the dataset in De Luca and Loperfido (2012), descriptive statistics and the results in Section 3 to assess the adequacy of the multivariate SGARCH model. The data are the percentage logarithmic daily returns (simply returns, henceforth) recorded from June 25, 2003 to June 23, 2008 in the French, Spanish and Danish stock markets. The full dataset, named All, is partitioned in two other datasets. The first (second) one, named Bad (Good), contains the returns in the three markets, that is the followers markets, recorded a day after that the US market, that is the leader market, showed a negative (positive) sign.

The measures of multivariate skewness described in Section 2 are very similar to each other, within each dataset (Table 1).

	Bad	**Good**	**All**
Mardia	1.743	0.687	0.256
Partial	1.697	0.662	0.258
Directional	1.718	0.607	0.239

Table 1: measures of multivariate skewness

By Theorem 4 in Section 3, this suggests that the corresponding third order cumulants are very similar to rank-one matrices, consistently with the SGARCH model with proportional feedback effects.

The first singular values of the third order cumulants correponding to the three data sets are much larger than the other ones (Table 2).

	Bad	**Good**	**All**
First	7.284	3.136	2.130
Second	0.005	0.025	0.020
Third	0.005	0.004	0.005

Table 2: singular values of third sample cumulants

Equivalently, all third sample cumulants are very well approximated by matrices of rank one. Again, this empirical finding is consistent with the SGARCH model with proportional feedback effects.

The first right singular vectors associated to the third sample cumulants of the three data sets are very similar to each other (Table 3).

	Bad	Good	All
First component	0.575	0.562	0.587
Second component	0.507	0.542	0.493
Third component	0.642	0.624	0.642

Table 3: First right singular vectors of the third sample cumulants

This fact and the previous ones imply that the third cumulants are nearly proportional to each others, consistently with the SGARCH model with proportional feedback effects. Inferential and computational issues are beyond the scope of this paper. However, we conjecture that sampling variability and numerical errors account for the small differences between model cumulants and sample cumulants.

Acknowledgements

The author is grateful to two anonymous referees whose comments greatly helped to improve the quality of the paper.

References

Chen, W.W. and Deo, R.S. (2004). Power transformations to induce normality and their applications. *J. R. Statist. Soc. B* **66**, 117–130.

Christiansen, M. and Loperfido, N. (2014). Improved approximation of the sum of random vectors by the skew-normal distribution. *J. App. Prob.*, to appear.

Davis, A.W. (1980). On the effects of moderate multivariate nonnormality on Wilks's likelihood ratio criterion. *Biometrika* **67**, 419-427.

De Lathauwer L., De Moor B. and Vandewalle V. (2001). Independent component analysis and (simultaneous) third-order tensor diagonalization. *IEEE Transactions on Signal Processing* **49**, 2262–2271.

De Luca, G. and Loperfido, N. (2012). Modelling multivariate skewness in financial returns: a SGARCH approach. *Eur. J. Fin.,* Digital Object Identifier: 10.1080/1351847X.2011.640342.

De Luca G., M. Genton M. and Loperfido N. (2006). A multivariate skew-GARCH model. *Adv. Econometrics* **20**, 33–57.

French K.R., Schwert W.G. and Stambaugh R.F. (1987). Expected stock returns and volatility. *J. Fin. Econ.* **19**, 3–29.

Gower, J.C. (2004). The geometry of biplot scaling. *Biometrika* **91**, 705-714.

Greenacre, M. and Hastie, T. (1987). The geometric interpretation of correspondence analysis. *J. Amer. Statist. Assoc.* **82**, 437-447.

Henze, N. (1997). Limit laws for multivariate skewness in the sense of Mòri, Rohatgi and Székely. *Statist. Prob. Lett.* **33**, 299-307.

Isogai, T. (1983). On measures of multivariate skewness and kurtosis. *Math. Japon.* **28**, 251-261.

Loperfido, N. (2011). Spectral analysis of the fourth moment matrix. *Lin. Alg. App.* **435**, 1837-1844.

Loperfido, N. (2013). Skewness and the linear discriminant function. *Statist. Prob. Lett.* **83**, 93-99.

Loperfido, N. and Guttorp, P. (2008). Network bias in air quality monitoring design. *Environmetrics* **19**, 661-671.

Loperfido N., Navarro J., Ruiz J.M. and Sandoval C.J. (2007). Some Relationships Between Skew-Normal Distributions and Order Statistics from Exchangeable Normal Random Vectors. *Comm. Statist.- Theory and Methods* **36**, 1719-1733.

Magnus, J.R. and Neudecker, H. (1979). The commutation matrix: some properties and applications. *Ann. Statist.* **7**, 381-394.

Malkovich, J.F. and Afifi, A.A. (1973). On Tests for multivariate normality. *J. Amer. Statist. Ass.* **68**, 176-179.

Mardia, K.V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika* **57**, 519-530.

Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979). *Multivariate Analysis.* London: Academic Press.

Martin, C.D. and Porter, A.M. (2012). The extraordinary SVD. *Amer. Math. Monthly* **119**, 838-851.

Mooijaart, A. (1985). Factor analysis for non-normal variables. *Psychometrika* **50**, 323-342.

Mòri T.F., Rohatgi V.K. and Székely G.J. (1993). On multivariate skewness and kurtosis. *Theory Probab. Appl.* **38**, 547-551.

Rao, C.R. and Rao, M.B. (1998). *Matrix Algebra and its Applications to Statistics and Econometrics.* World Scientific Co. Pte. Ltd., Singapore.

Van Hulle, M. M. (2005). Edgeworth approximation of multivariate differential entropy. *Neural Computation* **17** (9), 1903-1910.

Van Loan, C.F. and Pitsianis, N.P. (1993). Approximation with Kronecker products. In *Linear Algebra for large Scale and Real Time Applications*, M.S. Moonen and G.H. Golub, eds., 293-314, Kluwer publications.