Soft Computing (2025) 29:4247–4257
https://doi.org/10.1007/s00500-025-10600-4

FOCUS

An exterior algebra approach to generalised variances and cross-covariances

Henry P. Wynn[1] · Anatoly Zhigljavsky[2]

Accepted: 8 January 2025 / Published online: 28 May 2025

Abstract

It has been shown by Pronzato et al. (Bernoulli 23(4A):2617–2642, 2017; J Multivar Anal 168:276–289, 2018) that simplicial volumes formed by independent copies of random variables can be used to extend the definition of generalised variances. It is shown in this paper that exterior algebra is a natural environment in which to study these constructions. This is used to extend the formulation to covariances and correlations. The theory leads naturally to dispersion ordering, that is partial orderings in which one random variable is more disperse than another if one squared simplicial volume stochastically dominates the other.

Keywords Exterior algebra · Grassman algebra · Generalised covariance · Canonical correlation · Stochastic ordering · Dispersion ordering

1 Introduction

Despite its use, in areas such as information geometry, the role of multilinear algebra in statistical theory has been limited. However, as soon as determinants arise in some statistical context, particularly in multivariate analysis, one can claim that we are using multilinear algebra or multilinear geometry. This is true of previous work of the authors (Pronzato et al. 2017, 2018, 2019) which related the expected volume of random simplices, represented by determinants, to the determinants of covariance matrices and marginal covariance matrices; see also Gillard et al. (2022) where the technique of simplicial distances developed in Pronzato et al. (2017, 2018) has been used for detection of outliers and cluster analysis. The expected volumes of simplices have also played a part in definitions of dispersion orderings in previous work (Giovagnoli and Wynn 1995). The ideas can be traced back to the seminal work of Hotelling (1992) in canonical correlation analysis (CCA) and Wilks (Wilks 1932, 1960) in generalised variance. Results of Sect. 4 dealing with cross-covariances can be used in widening the interpretations of the techniques of the standard CCA as well as various extensions of CCA including the regularized CCA (Tenenhaus and Tenenhaus 2011) and deep CCA (Andrew et al. 2013). Note also an extensive use of cross-covariances in the methodology of time series analysis and forecasting called singular spectrum analysis, see Golyandina and Zhigljavsky (2013); Golyandina et al. (2018). The main aim of this paper is to promote the idea that exterior algebra is a natural environment in which to study and extend formulae and show that the inner product in exterior algebra is the key formula for our purposes.

We start with an elementary discussion. Thus, in statistics and probability theory variances and covariances are closely related to metrics. If X and Y are two jointly distributed one-dimensional random variables and $\mathbb{E}$ denotes expectation with respect to their joint distribution then

$$\mathbb{E}(|X-Y|^2) = \mathrm{var}(X) + \mathrm{var}(Y) - 2\,\mathrm{cov}(X,Y) + [\mathbb{E}(X) - \mathbb{E}(Y)]^2.$$

If X_1, X_2 are two independent copies of the random variable X then

$$\mathbb{E}(|X_1 - X_2|^2) = 2\ \mathrm{var}\ (X).$$

If X is a random vector with covariance matrix

$$C(X,X) = \mathbb{E}(XX^T) - \mathbb{E}(X)\mathbb{E}(X)^T,$$

Henry P. Wynn: Deceased on November 2, 2024.

✉ Anatoly Zhigljavsky
ZhigljavskyAA@cardiff.ac.uk

[1] London School of Economics, London, UK

[2] School of Mathematics, Cardiff University, Cardiff CF24 4AG, UK

then for Euclidean distance and i.i.d. copies X_1, X_2

$$\mathbb{E}||X_1 - X_2||^2 = 2\,\mathrm{trace}(C(X, X)). \tag{1}$$

The cross-covariance matrix between two random n-vectors, X and Y, is

$$C(X, Y) = \mathbb{E}(XY^T) - \mathbb{E}(X)\mathbb{E}(Y)^T.$$

In this case,

$$\mathrm{trace}(C(X, Y)) = \sum_{i=1}^{n} \mathrm{cov}(X_i, Y_i)$$

can be considered as an overall measure of covariance. The present paper revisits the authors' papers (Pronzato et al. 2017, 2018) with a straightforward use of the exterior product.

In the first part of the paper, we will consider standard vectors, that is, vectors extending from the origin so that simplices are formed with one vertex at the origin. But in the spirit of our previous work we return briefly in Sect. 5 to what we term *affine* simplices. For example, in one dimension this is the length of the line from point X_1 to point X_2 and a triangle in three or more dimensions is described by three points X_1, X_2, X_3, away from the origin. Sections 5 and 6 cover generalised covariances and cross-covariances and Sect. 7 discusses a natural application to dispersion orderings.

2 Exterior algebra

Our calculations are based on n-dimensional base vector space $\mathbb{R}^n$ over $\mathbb{R}$ with vectors $x^{(1)}, x^{(2)} \ldots$ written as column vectors

$$x = (x_1, \ldots, x_n)^T. \tag{2}$$

Looking forward to the next section we will write a random vector in $\mathbb{R}^n$ as $X = (X_1, \ldots, X_n)^T$ and use independent identically distributed random (vector) copies of a random n-vector X; similarly for Y.

We label the standard unit vectors in $\mathbb{R}^n$ as $e_1, \ldots, e_n$ so that we may express a vector $x \in \mathbb{R}^n$ as:

$$x = x_1 e_1 + \cdots + x_n e_n.$$

Note that any independent basis may be used, but the standard basis is easier conceptually. The book (Darling 1994) is an excellent introduction.

The outer product of two vectors $x, y \in \mathbb{R}^n$ is written $x \wedge y$. Starting with basis vectors we write formal expression which lie in a formal vector space $\bigwedge^2 \mathbb{R}^n$ whose basis vectors are all ordered pairs $e_i \wedge e_j, ; i, j = 1, \ldots, n;\ i < j$. Then, we have the decomposition

$$x \wedge y = \sum_{i<j} (x_i y_j - x_j y_i)\, e_i \wedge e_j.$$

The coefficients are the determinants of 2×2 matrices from the appropriate entries of x and y and are signed areas of the triangles formed by the corresponding 2-vectors and the origin.

Starting with the basis $\{e_1, \ldots, e_n\}$ of $\mathbb{R}^n$ the following rules uniquely define the wedge product. Given real scalars a, b and vectors x, y, z

1. $(ax + y) \wedge z = a(x \wedge y) + y \wedge z$,
2. $x \wedge (by + z) = b(x \wedge y) + x \wedge z$,
3. $x \wedge x = 0$.

We interpret the terms $e_i \wedge e_j$ as an abstract coding or place-holder of the two dimensional space spanned by e_i and e_j, but assigned an orientation expressed by a sign. From the above axioms it follows that $(x + y) \wedge (x + y) = x \wedge y + y \wedge x = 0$, so that

$$x \wedge y = -y \wedge x,$$

which shows the importance of signs.

The machinery extends to the space of high exterior powers $\bigwedge^p \mathbb{R}^n$ and we define the pth wedge product for vectors $x^{(1)}, \ldots, x^{(p)} \in \mathbb{R}^n$ by

$$x^{(i)} \wedge \cdots \wedge x^{(p)} = \sum_{i_1 < i_2 < < i_p} D_{i_1 < \cdots < i_p}\, e_{i_1} \wedge \cdots \wedge e_{i_p},$$

where $D_{i_1 < \cdots < i_p}$ is the determinant giving the p-dimensional volumes for directions coordinated by the terms: $e_{i_1} \wedge \cdots \wedge e_{i_p}$:

$$D_{i_1 < \cdots < i_p} = \det\,[x^{(i_1)} : \cdots : x^{(i_p)}].$$

A key construction for us is the inner product on $\bigwedge^p \mathbb{R}^n$. When $p = 2$, for $x \wedge y, u \wedge v$ in $\bigwedge^2 \mathbb{R}^n$ we define

$$\langle x \wedge y, u \wedge v \rangle = \det \begin{bmatrix} x^T u & x^T v \\ y^T u & y^T v \end{bmatrix}.$$

The inner product on $\bigwedge^p \mathbb{R}^n$ is defined as

$$\langle x^{(1)} \wedge \cdots \wedge x^{(p)}, y^{(1)} \wedge \cdots \wedge y^{(p)} \rangle = \det\{\langle x^{(i)}, y^{(j)} \rangle\},$$

where the inner product on the right hand side is the standard inner product. A matrix formulation is sometimes useful. Thus, the matrix $\langle x^{(i)}, y^{(j)} \rangle$ is $U^T V$ where $U = [x^{(1)} : \cdots x^{(p)}]$, $V = [y^{(1)} : \cdots : y^{(p)}]$.

In order to avoid too much notation we will refer to an ordered subset of size p, $i_1 < \cdots < i_p$, by J_p, being careful to fix the context. Thus $\sum_{J_p}$ is the summation over all $\binom{n}{p}$ ordered subsets J_p. This notation is also used to index marginal variables. Thus, x_{J_p} is the vector with entries, in order, $x_{i_j} : j = 1, \ldots, p$, and $\{X_{J_p}\}$ are the marginal random vectors corresponding to J_p.

3 Expectations, generalised variances and covariances

For the random version of x in (2) we write

$$X = X_1 e_1 + \cdots + X_n e_n.$$

If $\mathbb{E}$ denotes expectation with respect to the full joint distribution, then

$$\mathbb{E}(X) = \mathbb{E}(X_1)e_1 + \cdots + \mathbb{E}(X_n)e_n.$$

We assume that all random vectors have zero mean to make formulae a little easier to handle. Thus we define the covariance matrix of a random X as

$$C(X, X) = \mathbb{E}(XX^T),$$

the cross covariance between random vectors X and Y as

$$\mathrm{cov}(X, Y) = \mathbb{E}(XY^T),$$

and the full covariance matrix between X and Y as

$$C\left[\begin{pmatrix} X \\ Y \end{pmatrix}, \begin{pmatrix} X \\ Y \end{pmatrix}\right] = \begin{bmatrix} C(X,X) & C(X,Y) \\ C(X,Y)^T & C(Y,Y) \end{bmatrix}.$$

Definition 3.1 For two random variables X, Y with values in $\mathbb{R}^n$ define the generalised variances and the generalised covariance respectively by the following determinants: $\det(C(X, X))$, $\det(C(Y, Y))$ and $\det(C(X, Y))$.

These definitions will be used for marginal vectors X_{J_p}, Y_{J_p} in dimension p for all $1 \le p \le n$, so that we write, for example $\det(C_p(X, Y))$. The following is essentially similar to the result in Pronzato (1998), but with an alternative proof.

Lemma 3.2 *Let X and Y be two random p-vectors and let $X_{(1)}, \ldots X_{(p)}$ and $Y_{(1)}, \ldots Y_{(p)}$ be two sets of iid copies of X and Y, respectively. Then*

$$\mathbb{E}\left\{\det\left(\sum_{i=1}^{p} X^{(i)} Y^{(i)\top}\right)\right\} = p!\ \det(\mathrm{cov}(X, Y)).$$

Proof The Sylvester formula for an inverse of an invertible $p \times p$ matrix A is

$$A^{-1} = \mathrm{adj}(A)\det(A),$$

where the (i, j) entry of the adjugate $\mathrm{adj}(A)$, is, with appropriate sign, the determinant of the $(n-1) \times (n-1)$ cofactor formed by deleting rows i and j of A. If A is invertible and a, b are n-vectors we have the well known formula

$$\det(A + ab^T) = \det(A)(1 + b^T A^{-1} a).$$

We shall need the more general version which applies whether or not A necessarily invertible:

$$\det(A + ab^T) = \det(A) + b^T \mathrm{adj}(A)\, a.$$

The proof now proceeds by induction on p. The case $p = 1$ is immediate. Now,

$$\det\left(\sum_{i=1}^{p} X^{(i)} Y^{(i)\top}\right) = \det\left(\sum_{i=1}^{p-1} X^{(i)} Y^{(i)\top}\right) + Y_{(p)}^T \,\mathrm{adj}\left(\sum_{i=1}^{p-1} X^{(i)} Y^{(i)\top}\right) X_{(p)}.$$

The first term on the right is zero because the matrix does not have full rank. Then

$$\begin{aligned}
&\mathbb{E}\left\{\det\left(\sum_{i=1}^{p} X^{(i)} Y^{(i)\top}\right)\right\} \\
&= \mathbb{E}\left\{Y_{(p)}^T \,\mathrm{adj}\left(\sum_{i=1}^{p-1} X^{(i)} Y^{(i)\top}\right) X_{(p)}\right\} \\
&= \mathbb{E}\left\{\mathrm{trace}\left(Y_{(p)}^T \,\mathrm{adj}\left(\sum_{i=1}^{p-1} X^{(i)} Y^{(i)\top}\right) X_{(p)}\right)\right\} \\
&= \mathbb{E}\left\{\mathrm{trace}\left(\mathrm{adj}\left(\sum_{i=1}^{p-1} X^{(i)} Y^{(i)\top}\right) X_{(p)} Y_{(p)}^T\right)\right\} \\
&= \mathrm{trace}\left(\mathbb{E}\,\mathrm{adj}\left(\sum_{i=1}^{p-1} X^{(i)} Y^{(i)\top}\right) \mathbb{E}(X_{(p)} Y_{(p)}^T)\right) \quad (*) \\
&= \mathrm{trace}\left\{\mathbb{E}\left(\mathrm{adj}\left(\sum_{i=1}^{p-1} X^{(i)} Y^{(i)\top}\right) \mathbb{C}(X, Y)\right)\right\}.
\end{aligned}$$

where the transition (*) uses the independence between copies. Then, whether or not $C = C(X, Y)$ is invertible, the last formula reduces, by the property of adjugates, to:

$$(p-1)!\,\mathrm{trace}[\mathrm{adj}(C)C] = (p-1)!\,p\,\det(C) = p!\,\det(C),$$

as required. □.

Recall our notation for margins, namely that $X_{J_p}^{(1)}, X_{J_p}^{(2)}, \ldots, X_{J_p}^{(p)}$ and $Y_{J_p}^{(1)}, \ldots, Y_{J_p}^{(p)}$ are the J_p-margins of p independent copies of the n-vectors X and Y, respectively. Then, using the inner product in $\bigwedge^p \mathbb{R}^n$ we have the key lemma of the paper.

Lemma 3.3 *Let $(X^{(1)}, Y^{(1)}), \ldots, (X^{(p)}, Y^{(p)})$ be independent copies of the extended base vector (X, Y). Then*

$$\begin{aligned}
&\mathbb{E}(\langle X^{(1)} \wedge \cdots \wedge X^{(p)}, Y^{(1)} \wedge \cdots \wedge Y^{(p)}\rangle)\\
&\quad = \mathbb{E}\left\{\det\left(\{\langle X^{(i)}, Y^{(j)}\rangle\}\right)\right\}\\
&\quad = p! \sum_{J_p} \det\left(C_p(X_{J_p}, Y_{J_p})\right),
\end{aligned}$$

where the sum is over all (ordered) index sets $J_p \subset \mathbb{N}$ of size p.

Proof The first equality is from the definition of the inner product. The second follows by expanding $\det((C(X_J, Y_J))$ by the Binet–Cauchy theorem and applying Lemma 3.2 to every term. □.

Replacing Y by X in the two lemmas replaces all cross-covariances matrices by covariance matrices: $C(X, X)$, that is

$$\begin{aligned}
&\mathbb{E}(\langle X^{(1)} \wedge \cdots \wedge X^{(p)}, X^{(1)} \wedge \cdots \wedge X^{(p)}\rangle)\\
&\quad = \mathbb{E}\left\{\det\left(\{\langle X^{(i)}, X^{(j)}\rangle\}\right)\right\}\\
&\quad = p! \sum_{J_p} \det\left(C(X_{J_p}, X_{J_p})\right),
\end{aligned}$$

Note that before taking expectation the quantity

$$v_p(X) = \det(\{\langle X^{(i)}, X^{(j)}\rangle\})$$

is the volume of the p-dimensional simplex spanned by the $X^{(1)} \wedge \cdots \wedge X^{(p)}$, as studied in Pronzato et al. (2017). We thus have a decomposition of the expectation of the square of this volume and the covariances of the p-margins of the original random variable X. In the case of X, Y the wedge-product formula gives a new type of covariance based on product of the signed areas of two random simplices, one for X and one for Y.

4 Generalised cross-covariances and correlations

4.1 Definitions and a key property

As mentioned in the introduction, $\det(C(X, Y))$ considered as a generalised cross-covariance is not as well-known as Wilks's generalised variance $\det(C(X, X))$. Despite this we can proceed to the following definition derived from Lemma 3.3.

Definition 4.1 The generalised p-cross covariance of two random n-vectors X and Y is defined as

$$C_p(X, Y) = \sum_{J_p} \det(C(X_{J_p}, Y_{J_p})),$$

and the p-covariance for X (similarly, for Y) as

$$C_p(X, X) = \sum_{J_p} \det(C(X_{J_p}, X_{J_p})),$$

where the summation is over all ordered p-index sets J_p.

The only difference from the formula in Lemma 3.3 is the removal of the multiplier $p!$. Given the definitions of the p-generalised variances in Pronzato et al. (2017), we have the following natural definition:

Definition 4.2 The generalised p-correlation between random n-vectors X and Y is defined as

$$\mathrm{corr}_p(X, Y) = \frac{\sum_{J_p} \det(C(X_J, Y_J))}{\sqrt{\sum_{J_p} \det(C(X_{J_p}, X_{J_p})}\sqrt{\sum_{J_p} \det(C(Y_{J_p}, Y_{J_p}))}},$$

where the summations are over all ordered p-index sets J_p.

It is easily established that

$$-1 \le \mathrm{corr}_p(X, Y) \le 1$$

for all $2 \le p \le n$, by using the requirement that the joint covariance matrix of X and Y must be non-negative definite.

An interesting analysis arises in the full n-dimensional case when for random n-vectors $X, Y, C(X, X) = C(Y, Y) = I_n$, the $n \times n$ identity. We may arrive at this special case en route to computing canonical correlation, and we shall refer to this case as being canonical. Thus, using spectral square roots, if we take two random n-vectors U, V and set $X = C(U, U)^{-\frac{1}{2}} U$ and $Y = C(V, V)^{-\frac{1}{2}} V$ then $C(X, Y)$ is the canonical cross-correlation matrix and the covariance matrix for (X, Y) is

$$\begin{bmatrix} I_n & C(X, Y) \\ C(X, Y)^T & I_n \end{bmatrix}$$

This study of canonical correlation goes back to Hotelling (1992).

The fine structure of the relationship between X and Y can be studied via the cross-correlation matrix $C(X, Y)$. We have the following lemmas.

Lemma 4.3 *(i) For a n-vectors X, Y with $C(X, X) = C(Y, Y) = I_n$, $C = C(X, Y)$ is a valid cross-correlation matrix if and only if:*

$$CC^T \leq I_n,$$

where $\leq$ is the Loewner ordering, with equality if and only if:

$$\det(C) = 1,$$

which, in turn, holds if and only if

$$CC^T = I_n.$$

Proof $\det(CC^T) = 1$, so that all the eigenvalues of CCT are unity. This forces

$$(CC^T)^2 = C^T C$$

and $C^T C$ must be the identity projector. The converse is immediate. □

The condition $C(X, Y)^T C(X, Y) = I_n$ implies that $C(X, Y)$ is a rotation: formally a member of the orthogonal group $\mathrm{O}(n)$. So we have the informal statement that all extreme cross-correlations matrices, C, are related to rotations.

4.2 Two examples

Example 1 Let $n = 2$ and consider the covariance matrix in canonical form above. Then

$$C(X, Y) = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

If $\det(C(X, Y)) = 1$ then the general solution can be written

$$C(X, Y) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix},$$

for an angle $0 \leq \theta < \pi$.

In this case the set of $C(X, Y)$ is a representation of the rotation group, $O(2)$. For multiples of $\frac{\pi}{4}$, we have the subgroup which is the 16 dihedral order group, D_2, of permutation and sign changes with elements and representations:

$$\pm\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \pm\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \pm\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \pm\begin{bmatrix} & 0 & 1 \\ -1 & 1 \end{bmatrix},$$

$$\pm\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \pm\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},$$
$$\pm\frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}, \pm\frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Example 2 Take $n = 4$, again in canonical form, and

$$C(X, Y) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \bigotimes \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix},$$

so that $C(X, Y)$ is a member of $\mathrm{O}(2) \bigotimes \mathrm{O}(2)$. We compute:

$$\begin{aligned} \phi_1 &= \cos(s)\cos(t) \\ \phi_2 &= 1 - \tfrac{2}{3}(\sin(s)^2 + \sin(t)^2) \\ \phi_3 &= \cos(s)\cos(t) \\ \phi_4 &= 1\,. \end{aligned}$$

For example, if $s = 0$, $t = \pi/2$, then we have $\phi_1 = 0$, $\phi_2 = \frac{1}{2}$, $\phi_3 = 0$, $\phi_4 = 1$ and

$$C(X, Y) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix},$$

which is a member of $D_2 \bigotimes D_2$, as expected.

4.3 The eigenvalues of *C*

The eigenvalues of C may be complex, but condition $CC^T \leq I_n$ in Lemma (4.3) imposes restrictions.

Lemma 4.4 *For n-vectors X, Y with $C(X, X) = C(Y, Y) = I_n$ every eigenvalue λ of the cross-correlation matrix $C = C(X, Y))$ satisfies $|\lambda| \leq 1$.*

Proof We carry out the proof for the complex case. Let $z = u + iv$, with u and v real and $v \neq 0$, be the eigenvector corresponding to a λ. Then λ^*, the complex conjugate of λ, is the eigenvalue for the conjugate of z namely $z^* = u - iv$. Since $C^T z = \lambda z$ and $C^T z^* = \lambda^* z^*$

$$\begin{aligned} |\lambda|^2 \|z\|^2 &= \lambda^* \lambda z^{*\top} z \\ &= z^{*\top} CC^T z \\ &= u^T CC^T u + v^T CC^T v \;\; \text{(cross terms cancel)} \\ &\leq ||u||^2 + ||v||^2 = \|z\|^2, \end{aligned}$$

and cancelling $\|z\|^2$ gives the result. □

It is natural to ask whether in the canonical cross correlation case the matrix $C(X, Y)$ has a representation which might be thought of as a kind of PCA for cross correlations. This is indeed the case but since $C(X, Y)$ is not necessarily symmetric we need the Jordan form decomposition.

In the case that the eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ of $C(X, Y)$ are real and distinct there exists a matrix Q such that

$$Q\, C(X, Y)\, Q^{-1} = \mathrm{diag}(\{\lambda_1, \ldots, \lambda_n\}),$$

and if there are repeated roots then $Q\, C(X, Y)\, Q^{-1}$ has the usual Jordan block decomposition.

Complex eigenvalues occur in conjugate pairs: $\lambda = \lambda_j \pm \mu_j i$. For distinct conjugate pairs $QC(X, Y)Q^{-1}$ there is a version of the Jordan decomposition which gives 2×2 a blocks of the form

$$\begin{bmatrix} \lambda_j & -\mu_j \\ \mu_j & \lambda_j \end{bmatrix},$$

with extended forms when complex roots are repeated.

When the roots of $C(X, Y)$ are real we have the equivalent linear representation

$$Y_i = \lambda_i X_i,\ i = 1, \ldots, n.$$

But in the complex case we have for pairs $\{X_{j,1}, Y_{j,1}\}$, $\{X_{j,2}, Y_{j,2}\}$:

$$\begin{aligned} Y_{j,1} &= \lambda_j X_{j,1} - \mu_j X_{j,2}\,, \\ Y_{j,2} &= \mu_j X_{j,1} + \lambda_j X_{j,2}\,. \end{aligned}$$

Note, however, that the matrix $C^* = QC(X, Y)Q^{-1}$ is, in general, no longer the covariance between X and Y, but between QX and $Q^{-1}Y$. That is, by transforming $C(X, Y)$ to the canonical form may affect the canonical representation $C(X, X) = C(Y, Y) = I_n$.

In several fields this analysis is used to indicate the presence of *feedback*. Examples are in control theory and the closely related Granger causality in economics. We can, of course have a mixture of both real and complex eigenvalues.

5 The affine case

In Pronzato et al. (2017) the authors consider what we call here the affine case, motivated by (1). To aid explanation consider the first interesting example, namely triangles in three dimension.

Consider three i.i.d. copies $X^{(1)}, X^{(2)}, X^{(3)}$ in R^3, labelled as points A, B, C, respectively. They form a triangle ABC whose squared area is

$$\frac{1}{4}\, \det([X^{(1)} - X^{(3)} : X^{(2)} - X^{(3)}]^T \\ [X^{(1)} - X^{(3)} : X^{(2)} - X^{(3)}]).$$

In both cases we are considering the vectors from A to C and B to C. We can then expand by the Binet–Cauchy lemma and write the last expression as

$$\tfrac{1}{4}\, \big\{\det^2([X^{(1)} - X^{(3)} : X^{(2)} - X^{(3)}]_{12}) \\ + \det^2([X^{(1)} - X^{(3)} : X^{(2)} - X^{(3)}]_{13}) \\ + \det^2([X^{(1)} - X^{(3)} : X^{(2)} - X^{(3)}]_{23})\big\}\,.$$

This can be expressed using the wedge inner product as

$$\frac{1}{4}\, \langle (X^{(1)} - X^{(3)}) \wedge (X^{(2)} - X^{(3)}),\, (X^{(1)} - X^{(3)}) \wedge (X^{(2)} - X^{(3)}) \rangle.$$

It is natural to consider the covariance case, namely:

$$\langle (X^{(1)} - X^{(3)}) \wedge (X^{(2)} - X^{(3)}),\, (Y^{(1)} - Y^{(3)}) \wedge (Y^{(2)} - Y^{(3)}) \rangle,$$

the expansion of which is

$$\begin{aligned} &\det([X^{(1)} - X^{(3)} : X^{(2)} - X^{(3)}]_{12}) \\ &\quad \times \det([Y^{(1)} - Y^{(3)} : Y^{(2)} - Y^{(3)}]_{12}) \\ &+ \det([X^{(1)} - X^{(3)} : X^{(2)} - X^{(3)}]_{13}) \\ &\quad \times \det([Y^{(1)} - Y^{(3)} : Y^{(2)} - Y^{(3)}]_{13}) \\ &+ \det([X^{(1)} - X^{(3)} : X^{(2)} - X^{(3)}]_{23}) \\ &\quad \times \det([Y^{(1)} - Y^{(3)} : Y^{(2)} - Y^{(3)}]_{23})\,. \end{aligned}$$

Taking expectations we see that our generalised 2-covariance is the expectation of a sum of products of signed areas from blades of dimension $p = 2$. We then adapt the analysis of Sect. 3 to the affine case by extending with a vector of ones, $z = (1, 1, 1)^T$. Thus we replace vectors X by $\widetilde{X} = (X^T : z^T)^T$ and use the general version of the formulae

$$\det([X^{(1)} - X^{(3)} : X^{(2)} - X^{(3)}]_{12}) = \det([\widetilde{X^{(1)}_{12}} : \widetilde{X^{(2)}_{12}} : \widetilde{X^{(3)}_{12}}]).$$

Generalising the above argument, Lemma 3.3 is replaced by

Lemma 5.1 *Let $(X^{(1)}, Y^{(1)}), \ldots, (X^{(p+1)}, Y^{(p+1)})$ be independent copies of the base vector (X, Y). Then*

$$\begin{aligned} &E(\langle \widetilde{X^{(1)}} \wedge \cdots \wedge \widetilde{X^{(p+1)}}, \widetilde{Y^{(1)}} \wedge \cdots \wedge \widetilde{Y^{(p+1)}} \rangle) \\ &\quad = E\left\{\det\left\{\langle \widetilde{X^{(i)}}, \widetilde{Y^{(j)}} \rangle\right\}\right\} \\ &\quad = (p+1)! \sum_{J_p} \det\left(C(X_{J_p}, Y_{J_p})\right). \end{aligned}$$

When Y is replaced by X we obtain the main result in Pronzato et al. (2017). The results also extend in natural way to obtain an affine version of the development of the covariance representation in Sect. 4, with the analogous explanation in terms of the product of volumes of affine simplices.

6 Hodge star operator and the cross-covariance Pfaffian

The Hodge star construction, in the general case $(p, n-p)$, shows that for elements $X^{(1)} \wedge \cdots \wedge X^{(p)}$ in $\bigwedge^p$ and $Y^{(1)} \wedge \cdots \wedge Y^{(n-p)}$ in $\bigwedge^{n-p}$ there is a mapping, called the Hodge star operator, which takes $X^{(1)} \wedge \cdots \wedge X^{(p)}$ into its Hodge star dual, $\left(X^{(1)} \wedge \cdots \wedge X^{(p)}\right)^*$ in $\bigwedge^{n-p}$ such that

$$\begin{aligned} & X^{(1)} \wedge \cdots \wedge X^{(p)} \wedge Y^{(1)} \wedge \cdots \wedge Y^{(n-p)} \\ & = \langle \left(X^{(1)} \wedge \cdots \wedge X^{(p)}\right)^*, \\ & \quad Y^{(1)} \wedge \cdots \wedge Y^{(n-p)} \rangle \, e_1 \wedge \cdots e_n. \end{aligned}$$

We study the case $n = 2p$ so that $X^{(1)} \wedge \cdots \wedge X^{(p)}$ and $\left(X^{(1)} \wedge \cdots \wedge X^{(p)}\right)^*$ both have dimension p. Taking expectation and suppressing $e_1 \wedge \cdots \wedge e_n$ we have the identity

$$\mathrm{E}\left\{\det\left(X^{(1)} : \cdots : X^{(p)} : Y^{(1)} : \cdots : Y^{(p)}\right\}\right. \tag{3}$$

$$= \mathrm{E}\left\{\langle \left(X^{(1)} \wedge \cdots \wedge X^{(p)}\right)^*, Y^{(1)} \wedge \cdots \wedge Y^{(p)} \rangle\right\}. \tag{4}$$

Definition 6.1 Let $\left(X^1, Y^1), \cdots, (X^p, Y^p)\right)$ be independent i.i.d. copies of possibly correlated p-vectors with cross covariance C. Define $\phi(C)$, equivalently, by (3) or (4) above, as the (generalised) dual cross-covariance of C.

Expand in determinant form, so that:

$$(X^{(1)} \wedge \cdots \wedge X^{(p)}) = \sum_{i_1 < \cdots < i_p} D_{i_1 < \cdots < i_p} e_{i_1} \wedge \cdots \wedge e_{i_p}.$$

Then

$$(X^{(1)} \wedge \cdots \wedge X^{(p)})^* = \sum_{i_1 < \cdots < i_p} D_{i_1 < \cdots < i_p} (e_{i_1} \wedge \cdots \wedge e_{i_p})^*. \tag{5}$$

From the Hodge star theory the values of $(e_{i_1} \wedge \cdots \wedge e_{i_p})^*$ are all known. In summary, each $(e_{i_1} \wedge \cdots \wedge e_{i_p})^*$ is a particular complementary base element of $\bigwedge^p$ with an appropriate sign.

Then, rearranging (5) we transfer the star, again with appropriate sign, to $D_{i_1 < \cdots < i_p}$, and write

$$(X^{(1)} \wedge \cdots \wedge X^{(p)})^* = \sum_{i_1 < \cdots < i_p} D^*_{i_1 < \cdots < i_p} (e_{i_1} \wedge \cdots \wedge e_{i_p}). \tag{6}$$

We are now able to match terms in the Binet–Cauchy expansion in (6) and write

$$\begin{aligned} & \left\langle \left(X^{(1)} \wedge \cdots \wedge X^{(p)}\right)^*, Y^{(1)} \wedge \cdots \wedge Y^p \right\rangle \\ & = \sum_{i_1 < \cdots < i_p} D_{i_1 < \cdots < i_p} D^*_{i_1 < \cdots < i_p}. \end{aligned} \tag{7}$$

In particular, (7) gives a representation of $\phi(C)$ in terms of determinants of $p \times p$ covariance matrices, but with complementary index sets, rather than matched index sets as in Lemma 3.3.

Example 3 For $n = 2, \; p = 1$ and

$$C = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}$$

we have

$$\begin{aligned} \phi(C^*_{2,2}) &= E\left(\det\begin{bmatrix} X_1 & Y_1 \\ X_2 & Y_2 \end{bmatrix}\right) \\ &= E\,(X_1 Y_2 - X_2 Y_1) \\ &= c_{12} - c_{21}. \end{aligned}$$

Example 4 For $n = 4$, $p = 2$ and

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{bmatrix}$$

we obtain

$$\begin{aligned} \phi(C) &= E\left\{\det\begin{bmatrix} X_1^{(1)} & X_1^{(2)} & Y_1^{(1)} & Y_1^{(2)} \\ X_2^{(1)} & X_2^{(2)} & Y_2^{(1)} & Y_2^{(2)} \\ X_3^{(1)} & X_3^{(2)} & Y_3^{(1)} & Y_3^{(2)} \\ X_4^{(1)} & X_4^{(2)} & Y_4^{(1)} & Y_4^{(2)} \end{bmatrix}\right\} \\ &= -\det\begin{bmatrix} c_{13} & c_{14} \\ c_{23} & c_{24} \end{bmatrix} + \det\begin{bmatrix} c_{12} & c_{14} \\ c_{32} & c_{34} \end{bmatrix} + \det\begin{bmatrix} c_{12} & c_{13} \\ c_{42} & c_{43} \end{bmatrix}. \\ &\quad -\det\begin{bmatrix} c_{21} & c_{24} \\ c_{31} & c_{34} \end{bmatrix} + \det\begin{bmatrix} c_{21} & c_{23} \\ c_{41} & c_{43} \end{bmatrix} - \det\begin{bmatrix} c_{31} & c_{32} \\ c_{41} & c_{42} \end{bmatrix} \end{aligned}$$

It turns out that $\phi(C)$ is a recognisable quantity which is the subject of considerable research with many application in diverse fields, namely the *Pfaffian* of C, see Dress and Wenzel (1995).

The Pfaffian $\mathrm{pf}(A)$ of an antisymmetric square matrix ($A = -A^T$), is a special polynomial function of the entries of A, with integer coefficients, and with the property

$$[\mathrm{pf}(A)]^2 = \det(A).$$

In our case we set

$$A = C - C^T.$$

The following is the main result of this section, the proof of which can be developed using the arguments above, but which will be included in a subsequent more technical version.

Lemma 6.2 *If n is even, then the dual cross-covariance, $\phi(C)$, of the $n \times n$ cross-covariance matrix C is equal to the Pfaffian of the antisymmetric matrix $A = C - C^T$, and is the square root, with appropriate sign, of $det(A)$.*

Proof The following is a sketch. For n even, we first define a class of permutations Π that maps $\{1, 2, \ldots, n\}$ into blocks which consist of (disjoint) ordered pairs. For example, for $n = 4$ we may have $\{1, 2, 3, 4\} \to \{1, 4, 2, 3\}$, the pairs being $(1, 4)$ and $(2, 3)$. Let for $\pi \in \Pi$ the ordered pairs are $(i_1, j_1), (i_2, j_2), \ldots, (i_n, j_n)$. Then for any antisymmetric $n \times n$ matrix $A = \{a_{ij}\}$ with $n = 2p$ we have

$$\mathrm{pf}(A) = \sum_{\pi \in \Pi} \prod_{k=1}^{p} a_{(i_k, j_k)}. \tag{8}$$

We then use the fact that the p pairs $\big((X^1, Y^1), \cdots, (X^p, Y^p)\big)$ are independent i.i.d with mean zero. Many of the terms obtained by expanding the determinant in (7) are zero. Close inspection shows that the remaining terms give (8).

This representation shows that $\phi(C)$ is a function of the differences: $c_{ij} - c_{ji}$. In the case $n = 4$, $p = 2$ we have

$$\begin{aligned}\phi(C) = (c_{12} - c_{21})(c_{34} - c_{43}) - (c_{13} - c_{31})(c_{24} - c_{42}) \\ + (c_{23} - c_{32})(c_{14} - c_{41}).\end{aligned}$$

We can check this is equal to the determinant representation above.

This points to $\phi(C)$ being a rather special measure of the symmetry of C. The following is well known: for any real $n \times n$ antisymmetric matrix A there is an orthogonal matrix Q such that $B = QAQ^T$ has has the form of 2×2 antisymmetric blocks on the diagonal, but with zero diagonal:

$$\begin{bmatrix} 0 & \lambda_1 & 0 & 0 \ldots 0 \\ -\lambda_1 & 0 & 0 & 0 \ldots 0 \\ 0 & 0 & 0 & \lambda_2 \ldots 0 \\ 0 & 0 & -\lambda_2 & 0 \ldots 0 \\ \ldots & & & \\ 0 & 0 & & \ldots 0 \end{bmatrix}$$

In our case $A = C - C^T$ and $C = C(X, Y)$. In our earlier notation, we can consider

$$QCQ^T = C(QX, QY).$$

In addition,

$$\begin{aligned} QAQ^T &= QCQ^T - QC^TQ^T \\ &= C(QX, QY) - C(QX, QY)^T, \end{aligned}$$

which is the antisymmmetrized version of the the covariance $C' = C(X', Y')$ of variables $X' = QX,\ Y' = QY$. Let $C = \{c'_{ij}\}$. In this case $\phi(C) = \phi(C')$ and since $pf(B) = pf(A)$ we have

$$\phi(C) = \prod_{i=1\ldots,n-1} \left(c'_{i,i+1} - c'_{i+1,i}\right).$$

In summary, we can, after transformation, express $\phi(C)$ as a simple measure of symmetry.

It has been mentioned several times that the main concept in this and the authors' previous papers is to show that certain types of generalised variances and cross-covariances can be shown to be proportional to the expected volume, or squared volume, of random simplices. It should be pointed out, then, that the determinant in (7) is proportional to the (signed) volume of a random simplex in R^n formed by p random pairs $(X^{(j)}, Y^{(j)}, j = 1, \ldots, p$. From the properties of the Pfaffian this quantity is zero (for even n) if and only if the cross-covariance matrix, $C = C(X, Y)$ between X and Y is zero.

7 Stochastic dominance

Recall that standard stochastic dominance: $U \prec_{st} V$ is defined for univariate random variable U, V with cdf's $F_U(t), F_V(t)$ respectively if $F_V(x) \le F_U(x)$ for all $x \in \mathbb{R}$. Now, starting with the squared volume v_p of the p-dimensional spanned by the columns of an $n \times p$ matrix X there is a natural way to introduce a form of stochastic dominance, usually referred to as dispersion ordering. This is an extension of the version introduced in Giovagnoli and Wynn (1995) and studied by others eg Ayala and Lóópez-Díaz (2009).

Definition 7.1 For two random n-vectors $X^{(1)}$ and $X^{(2)}$ let X_1 and X_2 be the matrices whose columns are given by respectively p iid copies of $X^{(1)}$ and $X^{(2)}$. Then define $X^{(1)} \prec_p X^{(2)}$ if and only if

$$v_p(X^{(1)}) \prec_{st} v_p(X^{(2)})$$

Here we study the linear case by finding the class of $n \times n$ matrices A such that if

$$v_p(AZ) \leq v_p(Z)),$$

for all Z which (with abuse of notation), would immediately imply

$$AZ \prec_p Z,$$

for any random vector Z.

If X is the $n \times p$ matrix

$$X = [x^{(1)} : \ldots : x^{(p)}],$$

Then,

$$v_p(X) = \det(X^T X)$$

and

$$v_p(AX) = \det(X^T A^T AX).$$

So, we are required to find the class of $n \times n$ matrices A such that

$$\det(X^T A^T AX) \leq \det(X^T X),$$

for all $n \times p$ matrices X.

Let the SVD of $A^T A$ be

$$A^T A = P^T \Lambda P,$$

where $P^T P = I_n$, the $n \times n$ identity and $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is the vector of ordered eigen values $\lambda_1 \geq \cdots \geq \lambda_n$.

We replace X by $Y = P_X$, so that

$$\begin{aligned} X^T X &= Y^T Y \\ X^T A^T AX &= Y^T \Lambda Y \end{aligned}$$

The required conditions on A then reduce to conditions on the $\{\lambda_j\}$. We single out the $p \times p$ matrix holding the p largest eigenvalues: $\Lambda_1 = \operatorname{diag}(\lambda_1, \ldots, \lambda_p)$.

Theorem 7.2 *For an $n \times n$ matrix A*

$$v_p(AX) \leq v_p(X)$$

if and only if and only if $\det(\Lambda_1) \leq 1$.

Proof. Following the above working it is enough to show that $\det(Y^\top \Lambda Y) \leq \det(Y^\top Y)$ for all $p \times n$ matrices Y if and only if $\det(\Lambda_1) \leq 1$.

Split a $p \times n$ matrix Y into $p \times p$ matrix Y_1 and $p \times (n-p)$ matrix Y_2: $Y = (Y_1 : Y_2)$.

$$\begin{aligned} v_p(Y) &= \det(Y^\top Y) = \det(Y_1^\top Y_1 + Y_2^\top Y_2) \\ v_p(AY) &= \det(Y^\top \Lambda Y) = \det(Y_1^\top \Lambda_1 Y_1 + Y_2^\top \Lambda_2 Y_2) \end{aligned}$$

(i). Assume first that $\det(Y^\top \Lambda Y) \leq \det(Y^\top Y)$ for all $p \times n$ matrices Y. Choose Y_1 as the identity $p \times p$ matrix and Y_2 as an $p \times (n-p)$ matrix of zeros. Then $\det(Y^\top \Lambda Y) = \det(\Lambda_1)$ and $\det(Y^\top Y) = 1$.

(ii). Assume now that $\det(\Lambda_1) \leq 1$ and that Y_1 has full rank p. Expanding $d_A(Y)$ and $d(Y)$, we obtain

$$\begin{aligned} v_p(Y) &= \det(Y_1^\top \Lambda_1 Y_1) \\ &\qquad \det\left[I + \Lambda_2^{1/2} Y_2^\top (Y_1^\top \Lambda_1 Y_1)^{-1} Y_2 \Lambda_2^{1/2}\right], \\ v_p(AY) &= \det(Y_1^\top Y_1)\ \det\left[I + Y_2^\top (Y_1^\top Y_1)^{-1} Y_2\right]. \end{aligned}$$

As Y_1 and Λ are non-degenerate,

$$\begin{aligned} (Y_1^\top \Lambda_1 Y_1)^{-1} &= Y_1^{-1} \Lambda_1^{-1} (Y_1^\top)^{-1} \\ (Y_1^\top Y_1)^{-1} &= Y_1^{-1} (Y_1^\top)^{-1} \end{aligned}$$

This gives $Y_2^\top (Y_1^\top Y_1)^{-1} Y_2 = Z^\top Z$ where $Z = (Y_1^\top)^{-1} Y_2$ and

$$\begin{aligned} \Lambda_2^{1/2} Y_2^\top (Y_1^\top \Lambda_1 Y_1)^{-1} Y_2 \Lambda_2^{1/2} &= \Lambda_2^{1/2} Y_2^\top Y_1^{-1} \Lambda_1^{-1} (Y_1^\top)^{-1} Y_2 \Lambda_2^{1/2} \\ &= \Lambda_2^{1/2} Z^\top \Lambda_1^{-1} Z \Lambda_2^{1/2} \end{aligned}$$

As all diagonal elements of $\Lambda_2 = \operatorname{diag}(\lambda_{p+1}, \ldots, \lambda_n)$ are smaller than or equal to λ_p, we obtain

$$\det\left[I + Y_2 \Lambda_2^{1/2} Z^\top \Lambda_1^{-1} Z \Lambda_2^{1/2}\right] \leq \det\left[I + \lambda_p Z^\top \Lambda_1^{-1} Z\right]. \tag{9}$$

Moreover, all diagonal elements of Λ_1^{-1} are smaller than or equal to $1/\lambda_p$,

$$\Lambda_2^{1/2} Z^\top \Lambda_1^{-1} Z \Lambda_2^{1/2} \leq \lambda_p Z^\top \Lambda_1^{-1} Z \leq Z^\top Z$$

where these inequalities are valid in the Loewner sense.

Now since $\det(\Lambda_1) \leq 1$ and the matrices Y_1 and Λ are non-degenerate,

$$\det(Y_1^\top \Lambda_1 Y_1) \leq \det(Y_1^\top Y_1).$$

and from (9), we obtain

$$\begin{aligned} &\det\left[I + \Lambda_2^{1/2} Y_2^\top (Y_1^\top \Lambda_1 Y_1)^{-1} Y_2 \Lambda_2^{1/2}\right] \\ &\quad \leq \det\left[I + Y_2^\top (Y_1^\top Y_1)^{-1} Y_2\right] \end{aligned}$$

The last two inequalities imply that $\det(\Lambda_1) \leq 1$ imply the result.

Lemma 7.3 *Assume that* $B = diag(b_1, \ldots, b_n)$ *with* $0 \le b_i < 1$ *for all* i *and* C *is positive definite. Then*

$$\det(I + BCB) \le \det(I + C)$$

Proof. It is enough to calim that $\det(I + BCB)$ is monotonic as a function of each b_i. $\square$

Lemma 7.4 *Assume* $A = diag(a_1, \ldots, a_m)$ *with* $1 > a_1 \ge \cdots \ge a_m > 0$ *and* u *is a vector with all non-zero components.* $Auu^T A \le uu^T$ *if and only if* $A = cI$*. This inequality is true despite the fact that* $\Lambda_2^{1/2} Z^\top \Lambda_1^{-1} Z \Lambda_2^{1/2} \le \lambda_p Z^\top \Lambda_1^{-1} Z$*, is not true in general.*

Proof If $A = cI$ then clearly $Auu^\top A = c^2 uu^\top \le uu^\top$. Assume $Auu^\top A \le uu^\top$. Assume $a_q \ne a_{q+1}$ for some q. Choose x so that all components of x are 0 except for x_q and x_{q+1}. That is, we may assume $m = 2$ and $1 > a_1 > a_2 > 0$.

Now, consider

$$\det(Auu^\top A - uu^\top) = \det \begin{pmatrix} (1 - a_1^2)u_1^2 & (1 - a_1 a_2)u_1 u_2 \\ (1 - a_1 a_2)u_1 u_2 & (1 - a_2^2)u_2^2 \end{pmatrix} = -u_1^2 u_2^2 (a_1 - a_2)^2$$

Then, $\det(Auu^\top A - uu^\top) < 0$ unless $a_1 = a_2$ or either u_1 or u_2 is 0. $\square$

The development can be applied to the affine case, which was the case introduced in Giovagnoli and Wynn (1995) by replacing the vector X by $\tilde{X}$. That is to say A operates on $\tilde{X}$ and the simplices are affine simplices. This is an extended version of stochastic ordering defined in Giovagnoli and Wynn (1995) which corresponds to the case $p = 2$.

8 Conclusion

The expectation of the squared volume of random simplices formed by iid random vectors, is a natural generalisation of the expectation of squared length. In the latter case we obtain sums of variances (traces) and in the case of simplices the sums of the determinants of marginal covariance matrices. The expression in terms of determinants leads to a natural generalisation of Wilks's generalised variances. Exterior algebra gives a framework in which marginal determinants can be handled, in a sense simultaneously, via a generalized inner product. There are two special developments: generalised covariances/correlations and application to generalised dispersion orderings.

Acknowledgements Both authors are very grateful to Luc Pronzato for many useful comments. The authors are also grateful to Kristina Shubina (Crimtan) for technical assistance and IT-related consultations.

Author Contributions Equal.

Funding Self-funded.

Data availability Not applicable.

Code Availability Not applicable.

Declarations

Conflict of interest None.

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Materials availability Not applicable.

References

Andrew G, Arora R, Bilmes J, Livescu K (2013) Deep canonical correlation analysis. Proc Mach Learn Res 28(3):1247–1255

Ayala G, Lóópez-Díaz M (2009) The simplex dispersion ordering and its application to the evaluation of human corneal endothelia. J Multivar Anal 100(7):1447–1464

Darling RW (1994) Differential forms and connections. Cambridge University Press, Cambridge

Dress AW, Wenzel W (1995) A simple proof of an identity concerning Pfaffians of skew symmetric matrices. Adv Math 112(1):120–134

Gillard J, O'Riordan E, Zhigljavsky A (2022) Simplicial and minimal-variance distances in multivariate data analysis. J Stat Theory Pract 16(1):1–30

Giovagnoli A, Wynn HP (1995) Multivariate dispersion orderings. Stat Probab Lett 22(4):325–332

Golyandina N, Korobeynikov A, Zhigljavsky A (2018) Singular spectrum analysis with R. Springer, Berlin

Golyandina N, Zhigljavsky A (2013) Basic SSA. In: Singular spectrum analysis for time series, pp 11–70. Springer

Hotelling H (1992) Relations between two sets of variates. In: Breakthroughs in statistics (pp 162–190). Springer, New York

Oja H (1983) Descriptive statistics for multivariate distributions. Stat Probab Lett 1(6):327–332

Pronzato L (1998) On a property of the expected value of a determinant. Stat Probab Lett 39(2):161–165

Pronzato L, Wynn HP, Zhigljavsky A (2017) Extended generalised variances, with applications. Bernoulli 23(4A):2617–2642
Pronzato L, Wynn HP, Zhigljavsky A (2018) Simplicial variances, potentials and Mahalanobis distances. J Multivar Anal 168:276–289
Pronzato L, Wynn HP, Zhigljavsky A (2019) Bregman divergences based on optimal design criteria and simplicial measures of dispersion. Stat Pap 60(2):545–564
Sengupta A (2004) Generalized variance. Encycl Stat Sci 3(1)
Tenenhaus A, Tenenhaus M (2011) Regularized generalized canonical correlation analysis. Psychometrika 76:257–284
Wilks SS (1932) Certain generalizations in the analysis of variance. Biometrika 24(3/4):471–494
Wilks SS (1960) Multidimensional statistical scatter. Contributions to probability and statistics (Essays in Honor of Harold Hotelling) (Olkin, Ingram et al., eds.)

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