arXiv:1503.03212v2 [math.ST] 20 Jul 2015

Multivariate Generalized Gram-Charlier Series Using only Calculus of Several Variables, instead Tensor Calculus

Dharmani Bhaveshkumar C. DHARMANIBC@GMAIL.COM
Dhirubhai Ambani Institute of
Information & Communication Technology (DAIICT),
Gandhinagar, Gujarat, INDIA - 382001

Editor:

Abstract

The article derives multivariate Generalized Gram-Charlier (GGC) series that expands an unknown joint probability density function (PDF) of a random vector in terms of the differentiations of the joint PDF of a reference random vector. Though the higher order differentiations of a multivariate PDF result into multi-element arrays, the article uses only elementary calculus of several variables to achieve the extention of an existing derivation for GGC series in univariate to multivariate. The Tensor calculus is avoided by achieving the higher order differentiations of a multivariate PDF in vector notations through application of a specific Kronecker product based differentiation operator. The article also obtains expressions for the mutual relations between cumulants and moments of a random vector; integral form of a multivariate PDF; integral form of the multivariate Hermite polynomials; the multivariate Gram-Charlier A (GCA) series and others. The derived multivariate expressions are more elementary as using vector notations compare to coordinatewise tensor notations and more comprehensive as apparently more nearer to their counterparts for univariate. The approach has already been used to achieve the same advantages for deriving higher order cumulants of a random vector and multivariate Hermite polynomials.

Keywords: Multivariate Generalized Gram-Charlier series, Multivariate Gram-Charlier A series, Multivariate Hermite Polynomials, Kronecker Product

1. Introduction

The Generalized Gram-Charlier (GGC) series expands an unknown PDF in terms of the differentiations of a reference PDF, where the coefficients of expansion involve cumulant differences between those of an unknown PDF and a reference PDF. In general, the GC expansions are used for PDF and functions of PDF approximations. There have been used Poisson's distribution (Aroian, 1937), log-normal distribution (Aroian, 1947), binomial distribution (RIETZ, 1927), gamma distribution (Bowers, 1966) and others as reference PDFs for non-Gaussian density and related approximations. But, the GC expansion with Gaussian density as a reference PDF; for near-Gaussian density and related approximations; is the most popular and identified as the Gram-Charlier A (GCA) series. The rearrangement of the terms in GCA series results into the Edgeworth series. Though the Edgeworth series has better convergence property, the GCA series finds many applications. The GCA series is used for test of Gaussianity or near Gaussian PDF approximations (Aroian, 1947; Viehland, 1994; Girolami and Fyfe, 1996; Jondeau and Rockinger, 2001; Del Brio et al., 2009), for entropy

measure and independence measure approximations (Amari et al., 1996; Hyvärinen et al., 2001), for optima analysis through derivatives (Boscolo and Roychowdhury, 2003), for time-frequency analysis (Cohen, 1998) etc. towards Blind Source Separation (BSS) like Machine Learning applications ; as well, towards Economics, Chemistry, Astronomy and other application areas.

There exists multiple ways to derive univariate GCA series, as reported by Hald (2000). It's generalization to univariate GGC series is derived by Schleher (1977); Cohen (1998); Berberan-Santos (2007); Cohen (2011).

The multivariate GGC or GCA series derivation requires multivariate representations of - the Taylor series, the increasing order differentiations of a reference PDF and the cumulants. The higher order differentiations of a multivariate PDF and therefore the multivariate cumulants and multivariate Hermite polynomials require multilinear representations. It is known and acknowledged historically that going from matrix like notations to tensor notations for multivariate cumulants and multivariate Hermite polynomials have made the multivariate calculations more simpler and the representation more transparent (McCullagh, 1987). Along the same line, recently, it has been stated by (Terdik, 2002) that ' ... though the generalizations to higher dimensions may be considered straightforward, the methodology and the mathematical notations get quite cumbersome when dealing with the higher order derivatives of the characteristic function or other functions of a random vector ... '. Though the tensor notations have advantages over the matrix notations, they require componentwise separate representations and are more tedious compare to the vector notations. As a unique and comprehensive solution to this, there has been used an approach based only on elementary calculus of several variables (Terdik, 2002). The approach uses a specific Kronecker product based differential operator, identified as the 'K-derivative operator', to achieve vectorization of the Jacobian matrix of a multivariate vector function. The successive applications of this K-derivative operator achieves vectorization of the higher order derivatives also. Using this approach, there have been derived the Taylor series and the higher order cumulants by Terdik (2002); S. Rao Jammalamadaka and Terdik (May, 2006); as well, the multivariate Hermite polynomials by Terdik (2002); Holmquist (1996) in a more elementary way with more transparent representations.

There exists various derivations of multivariate GCA series using multi-element matrix notations for cumulants by Sauer and Heydt (1979); using recursive formula for Hermite polynomials by Berkowitz and Garner (1970); using tensor representation for cumulants and Hermite polynomials by (McCullagh, 1987, Chapter 5); using tensor representation for cumulants and involving multivariate Bell polynomials by Withers and Nadarajah (2014); using vector moments and vector Hermite polynomials by Holmquist (1996) and others. There exists various derivations for multivariate Edgeworth series (Davis, Dec., 1976; Skovgaard, 1986; Amari and Kumon, 1983; McCullagh, 1987; Kollo and von Rosen, 1998; Withers and Nadarajah, 2014). There is also derived generalized form of multivariate GC series, to represent a PDF through an another PDF, by (McCullagh, 1987, Chapter 5) in tensor notations. But, as per the author's knowledge, there exists neither the multivariate GGC series nor the multivariate Edgeworth series in vector notations.

Overall, to take advantages due to the recent advancement in representation, this article extends a specific derivation for univariate GGC series by Berberan-Santos (2007) to multivariate using only elementary calculus of several variables, instead of Tensor calculus.

As a by product, it also derives mutual relations between cumulants and moments; integral form of the multivariate PDF; integral form of the multivariate Hermite polynomials and the multivariate GCA series. All the derived multivariate expressions are more elementary as using vector notations and more comprehensive as apparently more nearer to their counterparts for univariate; compare to coordinatewise tensor notations. The theoretical results are verified using suitable examples.

Towards the aim of the article, the next Section 2 briefs some necessary background on the Kronecker product and a way to obtain vectorization of the higher order differentiations of a multivariate PDF. It also obtains the required multivariate Taylor series expansion using the derived notations. After the preliminary background, this article follows almost the same sequence for multivariate as that in (Berberan-Santos, 2007) for univariate. The Section 3 derives characteristic function, cumulants and moments of a random vector with their mutual relationships. The Section 4 obtains multivariate PDF in terms of its cumulants. The expressions for derivatives of multivariate Gaussian density and Hermite polynomials are derived in Section 5. The Section 6 derives multivariate GCA series by representing an unknown PDF in terms of the Gaussian PDF as a reference. The Section 7 derives GGC Expansion, representing an unknown PDF in terms of the a known reference PDF. The Section 8 derives an unknown characteristic function of a random vector in terms of a reference characteristic function. The Section 9 derives the same GGC expansion in a more compact way that summarizes the approach of the whole derivation. Finally, Section 10 concludes the article. For the sake of clarity, the calculation details and the proofs are kept in appendix at the end of the article.

2. Vectorization of the higher order differentiations

The section briefs the Kronecker product and the way it can be applied to achieve vectorization of the higher order differentiations of a multivariate PDF. Based on it, the multivariate Taylor series is obtained in vector notations. More details can be found on the Kronecker Product in (Magnus and Neudecker, 1999, Chapter 2), on achieving vectorization of the higher order differentiations in (Terdik, 2002; S. Rao Jammalamadaka and Terdik, May, 2006) and on the commutation matrices in (Magnus and Neudecker, 1999, Chapter 3, Section 7).

Definition 1 (Kronecker Product Operator $(\otimes)$) *The Kronecker Product Operator $(\otimes)$ between matrices $\mathbf{A}$ with size $p \times q$ and $\mathbf{B}$ with size $m \times n$ is defined as:*

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1q}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2q}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{p1}\mathbf{B} & a_{p2}\mathbf{B} & \cdots & a_{pq}\mathbf{B} \end{bmatrix} \tag{1}$$

The resultant matrix is of dimension $pm \times qn$.

Let $\mathbf{A}$ is with size $p \times 1$ and $\mathbf{B}$ is with size $m \times 1$, then $\mathbf{A} \otimes \mathbf{B}'$ is[1] a matrix with size $p \times m$. $\mathbf{A} \otimes \mathbf{A}$ is symbolically represented as $\mathbf{A}^{\otimes 2}$ and has size $p^2 \times 1$. In general, $\mathbf{A} \otimes \mathbf{A} \otimes \ldots \otimes \mathbf{A}$ (n times) is symbolically represented as $\mathbf{A}^{\otimes n}$ and has size $p^n \times 1$.

Definition 2 (Jacobian Matrix) *Let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_d)'$, $\boldsymbol{\lambda} \in \mathbb{R}^d$ and $\mathbf{f}(\boldsymbol{\lambda}) = (f_1(\lambda), f_2(\lambda), \ldots, f_m(\lambda))' \in \mathbb{R}^m$ be a differentiable m-component vector function. Then, Jacobian matrix of $\mathbf{f}(\boldsymbol{\lambda})$ ($\mathbf{J}(\mathbf{f})$) is an $m \times d$ matrix defined as under:*

$$\mathbf{J}(\mathbf{f}(\boldsymbol{\lambda})) = \frac{d\mathbf{f}}{d\boldsymbol{\lambda}} = \left[\frac{\partial \mathbf{f}}{\partial \lambda_1}, \frac{\partial \mathbf{f}}{\partial \lambda_2}, \ldots, \frac{\partial \mathbf{f}}{\partial \lambda_d}\right] = \begin{bmatrix} \frac{\partial f_1}{\partial \lambda_1} & \frac{\partial f_1}{\partial \lambda_2} & \cdots & \frac{\partial f_1}{\partial \lambda_d} \\ \frac{\partial f_2}{\partial \lambda_1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_m}{\partial \lambda_1} & \frac{\partial f_m}{\partial \lambda_2} & \cdots & \frac{\partial f_m}{\partial \lambda_d} \end{bmatrix} \tag{2}$$

Let the vector differential operator be defined as a column vector $\mathbf{D}_{\boldsymbol{\lambda}} = \left(\frac{\partial}{\partial \lambda_1}, \frac{\partial}{\partial \lambda_2}, \ldots, \frac{\partial}{\partial \lambda_d}\right)'$, then the Jacobian matrix, in terms of the $\mathbf{D}_{\boldsymbol{\lambda}}$, can be re-written as:

$$\mathbf{J}(\mathbf{f}(\boldsymbol{\lambda})) = \mathbf{D}_{\boldsymbol{\lambda}}(\mathbf{f}) = \mathbf{f}(\boldsymbol{\lambda})\mathbf{D}'_{\boldsymbol{\lambda}} = (f_1(\boldsymbol{\lambda}), f_2(\boldsymbol{\lambda}), \ldots, f_m(\boldsymbol{\lambda}))' \left(\frac{\partial}{\partial \lambda_1}, \frac{\partial}{\partial \lambda_2}, \ldots, \frac{\partial}{\partial \lambda_d}\right) \tag{3}$$

This implies that to match the definition of differentiation from matrix calculus, the vector differential operator should be applied from the right to the left. This is same as the requirement to be satisfied on generalization of vector derivative to matrix derivative as discussed by Magnus (2010). So, applying vector derivative operator from right to the left, has been kept as a rule throughout the article.

Definition 3 (The K-derivative Operator) *Let $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_d)'$, $\boldsymbol{\lambda} \in \mathbb{R}^d$, the vector differential operator $\mathbf{D}_{\boldsymbol{\lambda}} = \left(\frac{\partial}{\partial \lambda_1}, \frac{\partial}{\partial \lambda_2}, \ldots, \frac{\partial}{\partial \lambda_d}\right)'$ and $\mathbf{f}(\boldsymbol{\lambda}) = (f_1(\lambda), f_2(\lambda), \ldots, f_m(\lambda))' \in \mathbb{R}^m$ be a differentiable m-component vector function. Then, the Kronecker product between $\mathbf{D}_{\boldsymbol{\lambda}}$ and $\mathbf{f}(\boldsymbol{\lambda})$ is given as under:*

$$\mathbf{D}^{\otimes}_{\boldsymbol{\lambda}}\mathbf{f}(\boldsymbol{\lambda}) = \begin{bmatrix} f_1(\boldsymbol{\lambda}) \\ f_2(\boldsymbol{\lambda}) \\ \vdots \\ f_m(\boldsymbol{\lambda}) \end{bmatrix} \otimes \begin{bmatrix} \frac{\partial}{\partial \lambda_1} \\ \frac{\partial}{\partial \lambda_2} \\ \vdots \\ \frac{\partial}{\partial \lambda_d} \end{bmatrix} = Vec \begin{bmatrix} \frac{\partial f_1}{\partial \lambda_1} & \frac{\partial f_1}{\partial \lambda_2} & \cdots & \frac{\partial f_1}{\partial \lambda_d} \\ \frac{\partial f_2}{\partial \lambda_1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \frac{\partial f_m}{\partial \lambda_1} & \frac{\partial f_m}{\partial \lambda_2} & \cdots & \frac{\partial f_m}{\partial \lambda_d} \end{bmatrix}' \tag{4}$$

$$\Rightarrow \mathbf{D}^{\otimes}_{\boldsymbol{\lambda}}\mathbf{f}(\boldsymbol{\lambda}) = Vec\left(\frac{\partial \mathbf{f}}{\partial \boldsymbol{\lambda}'}\right)' = Vec\left(\frac{\partial}{\partial \boldsymbol{\lambda}}\mathbf{f}'\right) \tag{5}$$

where, the Vec operator converts $m \times d$ matrix into an $md \times 1$ column vector by stacking the columns one after an other. The operator $\mathbf{D}^{\otimes}_{\boldsymbol{\lambda}}$ is called Kronecker derivative operator or simply, K-derivative operator.

1. The symbol ' stands for Transpose of a matrix

Thus, the Kronecker product with the vector differential operator, obtains vectorization of the transposed Jacobian of a vector function. Corresponding to the definition, the k-th order differentiation is given by:

$$\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes k}\mathbf{f} = \mathbf{D}_{\boldsymbol{\lambda}}^{\otimes}\left(\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes k-1}\mathbf{f}\right) = [f_1(\lambda), f_2(\lambda), \ldots, f_m(\lambda)]' \otimes \left[\frac{\partial}{\partial\lambda_1}, \frac{\partial}{\partial\lambda_2}, \cdots, \frac{\partial}{\partial\lambda_d}\right]^{'\otimes k} \tag{6}$$

The $\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes k}\mathbf{f}$ is a column vector of dimension $md^k \times 1$. Some important properties of the K-derivative operator, those are useful in the further derivations, are listed in Appendix A.

2.1 Application of the K-derivative operator to the multivariate Taylor series

Let $\mathbf{x} = (X_1, X_2, ..., X_d)'$ be a d-dimensional column vector and $f(\mathbf{x})$ be the function of several variables differentiable in each variable. Using the defined K-derivative operator, the Taylor series for $f(\mathbf{x})$, expanding it at origin, is given as:

$$f(\mathbf{x}) = \sum_{m=0}^{m=\infty} \frac{1}{m!}\mathbf{c}(m,d)'\mathbf{x}^{\otimes m} \tag{7}$$

where, $\mathbf{c}(m,d)$ is the vector of dimension $d^m \times 1$ and given in terms of the derivative vector $\mathbf{D}_{\mathbf{x}} = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \cdots, \frac{\partial}{\partial x_d}\right)'$ as

$$\mathbf{c}(m,d) = \left(\mathbf{D}_{\mathbf{x}}^{\otimes m} f(\mathbf{x})\right)\Big|_{\mathbf{x}=\mathbf{0}} \tag{8}$$

The Taylor series expansion near $\boldsymbol{\lambda} = \mathbf{0}$, called the Maclaurian series, of some required functions based on the equation (7) are derived in appendix B.

3. Moments, cumulants and characteristic function of a random vector

Let $\mathbf{x} = (X_1, X_2, ..., X_d)'$ be a d-dimensional random vector and $f(\mathbf{x})$ be its joint PDF differentiable in each variable.

The Characteristic function ($\mathcal{F}$) of $\mathbf{x}$ is defined as the expected value of $\mathrm{e}^{i\mathbf{x}'\boldsymbol{\lambda}}$, where $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_d)'$, $\boldsymbol{\lambda} \in \mathbb{R}^d$. Also, both the characteristic function and the PDF are the Fourier Transform (F) of each other, in the sense they are dual.

$$\mathcal{F}_{\mathbf{x}}(\boldsymbol{\lambda}) = E\left\{\mathrm{e}^{i(\mathbf{x}'\boldsymbol{\lambda})}\right\} = \mathsf{F}(f(\mathbf{x})) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(\mathbf{x})\mathrm{e}^{i(\mathbf{x}'\boldsymbol{\lambda})}d\mathbf{x} \tag{9}$$

Expanding $\mathrm{e}^{i\mathbf{x}'\boldsymbol{\lambda}}$ using its Maclaurian series in equation (64) in appendix B, we get:

$$\mathcal{F}_{\mathbf{x}}(\boldsymbol{\lambda}) = \sum_{k=0}^{\infty}\mathbf{m}(k,d)'\frac{(i\boldsymbol{\lambda})^{\otimes k}}{k!} \tag{10}$$

where, $\mathbf{m}(k,d)$ is the k^{th} order moment vector of dimension $d^k \times 1$ and given by

$$\mathbf{m}(k,d) = \int_{\mathbb{R}^d} \mathbf{x}^{\otimes k} f(\mathbf{x}) d\mathbf{x}$$

$$\text{Also, } f(\mathbf{x}) = \mathsf{F}^{-1}(\mathcal{F}(\lambda)) = \mathsf{F}^{-1}\left(\sum_{k=0}^{\infty} \mathbf{m}(k,d)' \frac{(i\boldsymbol{\lambda})^{\otimes k}}{k!}\right) \tag{11}$$

$$= \sum_{k=0}^{\infty} \frac{\mathbf{m}(k,d)'}{k!} \left(\frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} (i\boldsymbol{\lambda})^{\otimes k} \mathrm{e}^{-i\mathbf{x}'\boldsymbol{\lambda}} d\boldsymbol{\lambda}\right)$$

$$= \sum_{k=0}^{\infty} (-1)^k \frac{\mathbf{m}(k,d)'}{k!} \mathbf{D}^{(k)} \delta(\mathbf{x}) \tag{12}$$

$$\left(\because \text{ Differentiation Property of FT: } \mathsf{F}\left(\mathbf{D}^{\otimes k} f(\mathbf{x})\right) = (i\boldsymbol{\lambda})^{\otimes k} \mathsf{F}(f(\mathbf{x}))\right)$$

The Moment Generating Function (MGF) of $f(\mathbf{X})$ is given as

$$\mathbf{M}(\boldsymbol{\lambda}) = E\left\{\mathrm{e}^{\mathbf{x}'\boldsymbol{\lambda}}\right\} = \int_{\mathbb{R}^d} f(\mathbf{X}) \mathrm{e}^{\mathbf{x}'\boldsymbol{\lambda}} d\mathbf{X} \tag{13}$$

$$= \sum_{k=0}^{\infty} \mathbf{m}(k,d)' \frac{\boldsymbol{\lambda}^{\otimes k}}{k!} \ (\because \text{ Expanding } \mathrm{e}^{\mathbf{x}'\boldsymbol{\lambda}}) \tag{14}$$

Assuming $\mathbf{M}(\boldsymbol{\lambda})$ and $\mathcal{F}(\boldsymbol{\lambda})$ are expanded using Taylor series,

$$\mathbf{m}(k,d) = \mathbf{D}_{\boldsymbol{\lambda}}^{\otimes k} \mathbf{M}(\boldsymbol{\lambda})\big|_{\boldsymbol{\lambda}=\mathbf{0}} = (-i)^k \mathbf{D}_{\boldsymbol{\lambda}}^{\otimes k} \mathcal{F}_{\mathbf{x}}(\boldsymbol{\lambda})\big|_{\boldsymbol{\lambda}=\mathbf{0}} \tag{15}$$

The Cumulant Generating Function (CGF) of $f(\mathbf{X})$ is given by,

$$\mathbf{C}(\boldsymbol{\lambda}) = \ln \mathbf{M}(\boldsymbol{\lambda}) = \sum_{k=1}^{\infty} \mathbf{c}(k,d)' \frac{\boldsymbol{\lambda}^{\otimes k}}{k!} \tag{16}$$

where, $\mathbf{c}(k,d)$ is the k^{th} order cumulant vector of dimension $d^k \times 1$.
The Cumulant Generating Function (CGF) of $f(\mathbf{X})$ can also be defined using the characteristic function, as under:

$$\mathcal{C}(\boldsymbol{\lambda}) = \ln \mathcal{F}(\boldsymbol{\lambda}) = \sum_{k=1}^{\infty} \mathbf{c}(k,d)' \frac{(i\boldsymbol{\lambda})^{\otimes k}}{k!} \tag{17}$$

Assuming $\mathbf{C}(\boldsymbol{\lambda})$ and $\mathcal{C}(\boldsymbol{\lambda})$ have been expanded using Taylor series,

$$\mathbf{c}(k,d) = \mathbf{D}_{\boldsymbol{\lambda}}^{\otimes k} \mathbf{C}(\boldsymbol{\lambda})\big|_{\boldsymbol{\lambda}=\mathbf{0}} = (-i)^k \mathbf{D}_{\boldsymbol{\lambda}}^{\otimes k} \mathcal{C}_{\mathbf{x}}(\boldsymbol{\lambda})\big|_{\boldsymbol{\lambda}=\mathbf{0}} \tag{18}$$

3.1 Relation between the cumulant vectors and the moment vectors

The relation between the moments and the cumulants is given by combining equation (14) and equation (16) as below:

$$\mathbf{M}(\boldsymbol{\lambda}) = \sum_{k=0}^{\infty} \mathbf{m}(k,d)' \frac{\boldsymbol{\lambda}^{\otimes k}}{k!} = \exp\left(\sum_{k=1}^{\infty} \mathbf{c}(k,d)' \frac{\boldsymbol{\lambda}^{\otimes k}}{k!}\right) \tag{19}$$

For $k = 1$, using equation (15), we get:

$$\mathbf{m}(1,d) = \mathbf{D}_{\boldsymbol{\lambda}}^{\otimes 1}\mathbf{M}(\boldsymbol{\lambda})\Big|_{\boldsymbol{\lambda}=\mathbf{0}}$$

Applying K-derivative ($\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes}$) to equation (19),

$$\sum_{p=1}^{\infty}\mathbf{m}(p,d)'\frac{\left(\boldsymbol{\lambda}^{\otimes(p-1)}\otimes\mathbf{I}_d\right)}{(p-1)!}\Bigg|_{\boldsymbol{\lambda}=\mathbf{0}} = \sum_{p=1}^{\infty}\mathbf{c}(p,d)'\frac{\left(\boldsymbol{\lambda}^{\otimes(p-1)}\otimes\mathbf{I}_d\right)}{(p-1)!}\otimes\exp\left(\sum_{q=1}^{\infty}\mathbf{c}(q,d)'\frac{\boldsymbol{\lambda}^{\otimes q}}{q!}\right)\Bigg|_{\boldsymbol{\lambda}=\mathbf{0}}$$
$$\Rightarrow \mathbf{m}(1,d) = \mathbf{c}(1,d)$$

Similarly, based on equation (15), for any k taking k^{th} order K-derivative of above equation (19) on both sides relates $\mathbf{m}(k,d)$ with $\mathbf{c}(k,d)$. For example, the cases for $k = 2$ and $k = 3$ are shown in Appendix C.

Overall, the moments in terms of the cumulants can be summarized as under:

$$\begin{aligned}
\mathbf{m}(1,d) &= \mathbf{c}(1,d)\\
\mathbf{m}(2,d) &= \mathbf{c}(2,d) + \mathbf{c}(1,d)^{\otimes 2}\\
\mathbf{m}(3,d) &= \mathbf{c}(3,d) + 3\mathbf{c}(2,d)\otimes\mathbf{c}(1,d) + \mathbf{c}(1,d)^{\otimes 3}\\
\mathbf{m}(4,d) &= \mathbf{c}(4,d) + 4\mathbf{c}(3,d)\otimes\mathbf{c}(1,d) + 3\mathbf{c}(2,d)^{\otimes 2} + 6\mathbf{c}(2,d)\otimes\mathbf{c}(1,d)^{\otimes 2} + \mathbf{c}(1,d)^{\otimes 4}\\
\mathbf{m}(5,d) &= \mathbf{c}(5,d) + 5\mathbf{c}(4,d)\otimes\mathbf{c}(1,d) + 10\mathbf{c}(3,d)\otimes\mathbf{c}(2,d) + 10\mathbf{c}(3,d)\otimes\mathbf{c}(1,d)^{\otimes 2}\\
&\quad + 15\mathbf{c}(2,d)^{\otimes 2}\otimes\mathbf{c}(1,d) + 10\mathbf{c}(2,d)\otimes\mathbf{c}(1,d)^{\otimes 3} + \mathbf{c}(1,d)^{\otimes 5}\\
\mathbf{m}(6,d) &= \mathbf{c}(6,d) + 6\mathbf{c}(5,d)\otimes\mathbf{c}(1,d) + 15\mathbf{c}(4,d)\otimes\mathbf{c}(2,d) + 15\mathbf{c}(4,d)\otimes\mathbf{c}(1,d)^{\otimes 2} + 10\mathbf{c}(3,d)^{\otimes 2}\\
&\quad + 60\mathbf{c}(3,d)\otimes\mathbf{c}(2,d)\otimes\mathbf{c}(1,d) + 20\mathbf{c}(3,d)\otimes\mathbf{c}(1,d)^{\otimes 2} + 15\mathbf{c}(2,d)^{\otimes 3}\\
&\quad + 45\mathbf{c}(2,d)^{\otimes 2}\otimes\mathbf{c}(1,d)^{\otimes 2} + 15\mathbf{c}(2,d)\otimes\mathbf{c}(1,d)^{\otimes 4} + \mathbf{c}(1,d)^{\otimes 6}
\end{aligned} \tag{20}$$

The above equations can be represented through more compact formulas as under. The equation (21) gives generalized k^{th} order d-variate cumulant vector in terms of the moment vectors and the equation (22) gives the cumulant vectors in terms of the moment vectors.

$$\mathbf{m}(k,d) = \sum_{p=0}^{k-1}\binom{k-1}{p}\mathbf{K}_{\mathfrak{p}h\leftrightarrow l}^{-1}\mathbf{c}(k-p,d)\otimes\mathbf{m}(p,d) \tag{21}$$

$$\mathbf{c}(k,d) = \mathbf{m}(k,d) - \sum_{p=1}^{k-1}\binom{k-1}{p}\mathbf{K}_{\mathfrak{p}h\leftrightarrow l}^{-1}\mathbf{c}(k-p,d)\otimes\mathbf{m}(p,d) \tag{22}$$

where, $\mathbf{K}_{\mathfrak{p}h\leftrightarrow l}^{-1}$ is a specific commutation matrix with corresponding dimensions that changes the order of the cumulants for Kronecker product such that the expression has decreasing order cumulants from left to the right, i.e. the higher order cumulant vector on the left and the lower order cumulant vector on the right. The derived multivariate expressions reduce to the following expressions for dimension $d = 1$ and are exactly same as those derived in

(Berberan-Santos, 2007).

$$\begin{aligned} c(k,1) &= m(k,1) - \sum_{p=1}^{k-1} \binom{k-1}{p} c(k-p,1) m(p,1) \\ \text{or more simply, } c_k &= m_k - \sum_{p=1}^{k-1} \binom{k-1}{p} c_{k-p} m_p \end{aligned} \tag{23}$$

Thus, the derived multivariate expressions are elementary vector extensions to those for univariate.

4. Multivariate PDF representation in terms of the cumulants

From equation (11) and equation (17), the multivariate PDF $f(\mathbf{x})$ can be written as:

$$f(\mathbf{x}) = \mathsf{F}^{-1}(\mathrm{e}^{\mathcal{C}(\boldsymbol{\lambda})}) = \left(\frac{1}{2\pi}\right)^d \int_{\mathbb{R}^d} \exp\left(\sum_{k=1}^{\infty} \mathbf{c}(k,d)' \frac{(i\boldsymbol{\lambda})^{\otimes k}}{k!}\right) \exp\left(-i\mathbf{x}'\boldsymbol{\lambda}\right) d\boldsymbol{\lambda} \tag{24}$$

As PDF is a real function and $\mathrm{Re}(e^{A+iB}e^{-iC}) = e^A \cos(B-C) = e^A \cos(C-B)$, the equation (24) can be re-written as:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \exp\left(\sum_{k=1}^{\infty} \frac{\mathbf{c}(2k,d)'}{2k!} (i\boldsymbol{\lambda})^{\otimes 2k}\right) \cos\left(\mathbf{x}'\boldsymbol{\lambda} + \sum_{k=1}^{\infty} \frac{\mathbf{c}(2k-1,d)'}{(2k-1)!} (i)^{2k} (\boldsymbol{\lambda})^{\otimes(2k-1)}\right) d\boldsymbol{\lambda} \tag{25}$$

The integrand in this equation is an even function. So,

$$f(\mathbf{x}) = \frac{1}{(\pi)^d} \int_{(\mathbb{R}^+)^d} \exp\left(\sum_{k=1}^{\infty} \frac{\mathbf{c}(2k,d)'}{2k!} (i\boldsymbol{\lambda})^{\otimes 2k}\right) \cos\left(\mathbf{x}'\boldsymbol{\lambda} + \sum_{k=1}^{\infty} \frac{\mathbf{c}(2k-1,d)'}{(2k-1)!} (i)^{2k} (\boldsymbol{\lambda})^{\otimes(2k-1)}\right) d\boldsymbol{\lambda} \tag{26}$$

where, $\mathbb{R}^+ = \{x \in \mathbb{R} : x \geq 0\}$. The equation (25) and the equation (26) give a multivariate PDF in terms of the cumulants. As they are derived using Taylor series expansion, the infinite order differentiability is an implicit assumption. The equations can be verified using known PDF examples with finite number of moments, let say the impulse delta density function and the Gaussian density function. The impulse delta density function has only first moment and all other moments are zero. Using this knowledge in equation (25),

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \cos\left(\mathbf{x}'\boldsymbol{\lambda} - \mathbf{c}(1,d)'\boldsymbol{\lambda}\right) d\boldsymbol{\lambda} \qquad (27) \\ &= \delta(\mathbf{x} - \mathbf{c}(1,d)) \text{ (shifted impulse delta function)} \end{aligned}$$

The Gaussian density function has first two moments and all other moments are zero. Using this knowledge in equation (25),

$$\begin{aligned} G(\mathbf{x}) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \exp\left(-\frac{\mathbf{c}(2,d)'}{2} \boldsymbol{\lambda}^{\otimes 2}\right) \cos\left(\mathbf{x}'\boldsymbol{\lambda} - \mathbf{c}(1,d)'\boldsymbol{\lambda}\right) d\boldsymbol{\lambda} \qquad (28) \\ &= (2\pi)^{-d/2} \left|\mathbf{C}_{\mathbf{x}}\right|^{-1/2} \exp\left(-\left(inv\ \mathbf{c}(2,d)\right)' \left(\mathbf{x} - \mathbf{c}(1,d)\right)^{\otimes 2}\right) \qquad (29) \\ &\quad (\text{ See Appendix D for proof}) \end{aligned}$$

where, $inv\ \mathbf{c}(2,d) = Vec\left(\mathbf{C}_{\mathbf{x}}^{-1}\right)$.

5. The multivariate Hermite polynomials in integral form

An interesting application of the integral form of multivariate PDF representation is achieved in this section. The multivariate Gaussian expressed as in equation (28) is used to derive it's differentials and Hermite polynomials in a simple way. Taking k^{th} K-derivative of $G(\mathbf{x})$,

$$G^{(k)}(\mathbf{x}) = \mathbf{D}_{\mathbf{x}}^{\otimes k} G(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \boldsymbol{\lambda}^{\otimes k} \exp\left(-\frac{\mathbf{c}(2,d)'}{2}\boldsymbol{\lambda}^{\otimes 2}\right) \cos\left(\left(\mathbf{x} - \mathbf{c}(1,d)\right)'\boldsymbol{\lambda} + \frac{k\pi}{2}\right) d\boldsymbol{\lambda} \tag{30}$$

Based on this, the multivariate Hermite polynomials defined by Holmquist (1996) are derived as under:

$$\mathbf{H}_k(\mathbf{x}; \mathbf{0}, \mathbf{C}_{\mathbf{x}}) = [G(\mathbf{x}; \mathbf{0}, \mathbf{C}_{\mathbf{x}})]^{-1} (-1)^k \left(\mathbf{C}_{\mathbf{x}} \mathbf{D}_x\right)^{\otimes k} G(\mathbf{x}; \mathbf{0}, \mathbf{C}_{\mathbf{x}}) \tag{31}$$

$$\begin{aligned} \text{where, } G(\mathbf{x}; \mathbf{0}, \mathbf{C}_{\mathbf{x}}) &= |\mathbf{C}_{\mathbf{x}}|^{-1/2} (2\pi)^{-d/2} \exp\left(-\frac{1}{2}\mathbf{x}'\mathbf{C}_{\mathbf{x}}^{-1}\mathbf{x}\right) \\ &= |\mathbf{C}_{\mathbf{x}}|^{-1/2} (2\pi)^{-d/2} \exp\left(-\frac{1}{2}\left(Vec\ \mathbf{C}_{\mathbf{x}}^{-1}\right)' \mathbf{x}^{\otimes 2}\right) \end{aligned} \tag{32}$$

This is equivalent to the 1-dimensional definition of Hermite polynomials[2] by Rodrigues's formula, except the introduction of matrix $\mathbf{C}_{\mathbf{x}}$.

$$H_k(x) = [G(x)]^{-1}(-1)^k \frac{d^k}{dx^k} G(x) \text{ where, } G(x) = \frac{1}{2\pi} \mathrm{e}^{-\frac{1}{2}x^2} \tag{33}$$

$$\begin{aligned} \mathbf{H}_k(\mathbf{x}; \mathbf{0}, \mathbf{C}_{\mathbf{x}}) = (2\pi)^{d/2} |\mathbf{C}_{\mathbf{x}}|^{1/2} (-1)^k (\mathbf{C}_{\mathbf{x}})^{\otimes k} \exp\left(\frac{1}{2}\left(Vec\ \mathbf{C}_{\mathbf{x}}^{-1}\right)' \mathbf{x}^{\otimes 2}\right) \\ \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \boldsymbol{\lambda}^{\otimes k} \exp\left(-\frac{\mathbf{c}(2,d)'}{2}\boldsymbol{\lambda}^{\otimes 2}\right) \cos\left(\left(\mathbf{x} - \mathbf{c}(1,d)\right)'\boldsymbol{\lambda} + \frac{k\pi}{2}\right) d\boldsymbol{\lambda} \end{aligned} \tag{34}$$

Taking $\mathbf{C}_{\mathbf{x}} = \mathbf{I}_d$, where $\mathbf{I}_d$ is $d \times d$ identity matrix; using the property $(Vec\ \mathbf{I}_d)'\mathbf{x}^{\otimes 2} = \mathbf{x}'\mathbf{x}$ and using change of variable as $\boldsymbol{\lambda}/\sqrt{2} = \mathbf{u}$ - the integral form of multivariate Hermite polynomials is obtained as under:

$$\mathbf{H}_k(\mathbf{x}; \mathbf{0}, \mathbf{I}_d) = (2)^{\frac{k+d+1}{2}} (\pi)^{-d/2} \exp\left(\frac{1}{2}\mathbf{x}'\mathbf{x}\right) \int_{\mathbb{R}^d} \mathbf{u}^{\otimes k} \exp\left(-\mathbf{u}'\mathbf{u}\right) \cos\left(\sqrt{2}\mathbf{x}'\mathbf{u} - \frac{k\pi}{2}\right) d\mathbf{u} \tag{35}$$

The result in equation (30) can also be obtained applying the derivative property of Fourier transform (F). The k^{th} derivative of $G(x)$ is given by,

$$G^{(k)}(\mathbf{x}) = \frac{(-1)^k}{(2\pi)^d} \int_{\mathbb{R}^d} (i\boldsymbol{\lambda})^{\otimes k} \mathsf{F}(G(\mathbf{x})) \mathrm{e}^{-i\mathbf{x}'\boldsymbol{\lambda}} d\boldsymbol{\lambda} \tag{36}$$

2. This is the 'probabilists' Hermite polynomials and not the 'physicists' Hermite polynomials used by Berberan-Santos (2007).

As PDF is a real function,

$$G^{(k)}(\mathbf{x}) = \frac{(-1)^k}{(2\pi)^d} \int_{\mathbb{R}^d} \boldsymbol{\lambda}^{\otimes k} \mathrm{Re}\left[i^k \mathsf{F}(G(\mathbf{x})) \mathrm{e}^{-i\mathbf{x}'\boldsymbol{\lambda}} \right] d\boldsymbol{\lambda} \tag{37}$$

Comparing equation (30) with the identity of $G(\mathbf{x})$ as $G(\mathbf{x}) = \mathrm{Re}\left[\mathsf{F}^{-1}\left(\mathsf{F}(G(\mathbf{x}))\right)\right]$, the $\mathsf{F}(G(\mathbf{x}))$ is obtained as under:

$$\mathsf{F}(G(\mathbf{x})) = \exp\left(-\frac{\mathbf{c}(2,d)'}{2} \boldsymbol{\lambda}^{\otimes 2} \right) \tag{38}$$

This achieves the differentiation of $G(\mathbf{x})$ in equation (37), same as in equation (30), through a different approach.

6. Multivariate Gram-Charlier A series

Till now, the article has derived - an unknown PDF expressed in terms of its cumulants in equation (25); the Gaussian density function expressed in terms of its cumulants in equation (28) and the Hermite polynomials in equation (35). Based on them, the multivariate Gram Charlier A series expansion that expresses the unknown PDF using Gaussian density as a reference can be obtained. The expansion assumes first and second order cumulants being same for both the unknown PDF and the reference PDF. Using the expansion $\exp(A+B)\cos(C+D) = \exp(A)\exp(B)(\cos C \cos D - \sin C \sin D)$, the equation (25) can be re-written as:

$$\begin{aligned} f(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} & \exp\left(-\frac{\mathbf{c}(2,d)'}{2} \boldsymbol{\lambda}^{\otimes 2} \right) \exp\left(\sum_{k=2}^{\infty} \frac{\mathbf{c}(2k,d)'}{2k!} (i\boldsymbol{\lambda})^{\otimes 2k} \right) \\ & \left\{ \cos((\mathbf{x} - \mathbf{c}(1,d))'\boldsymbol{\lambda}) \cos\left(\sum_{k=2}^{\infty} \frac{\mathbf{c}(2k-1,d)'}{(2k-1)!} (i)^{2k} (\boldsymbol{\lambda})^{\otimes(2k-1)} \right) \right. \\ & \left. - \sin((\mathbf{x} - \mathbf{c}(1,d))'\boldsymbol{\lambda}) \sin\left(\sum_{k=2}^{\infty} \frac{\mathbf{c}(2k-1,d)'}{(2k-1)!} (i)^{2k} (\boldsymbol{\lambda})^{\otimes(2k-1)} \right) \right\} d\boldsymbol{\lambda} \end{aligned} \tag{39}$$

Using the expansions in Appendix B, parts of the equation (39) can be approximated upto maximum 6^{th}-order statistics as under:

$$\begin{aligned} & \exp\left(\sum_{k=2}^{\infty} \frac{\mathbf{c}(2k,d)'}{2k!} (i\boldsymbol{\lambda})^{\otimes 2k} \right) \cos\left(\sum_{k=2}^{\infty} \frac{\mathbf{c}(2k-1,d)'}{(2k-1)!} (i)^{2k} (\boldsymbol{\lambda})^{\otimes(2k-1)} \right) = \\ & 1 + \left(\frac{\mathbf{c}(4,d)'\boldsymbol{\lambda}^{\otimes 4}}{4!} - \frac{\mathbf{c}(6,d)'\boldsymbol{\lambda}^{\otimes 6}}{6!} \right) - \frac{1}{2} \left(\frac{\mathbf{c}(3,d)'\boldsymbol{\lambda}^{\otimes 3}}{3!} \right)^{\otimes 2} + \dots \end{aligned} \tag{40}$$

$$\begin{aligned} & \exp\left(\sum_{k=2}^{\infty} \frac{\mathbf{c}(2k,d)'}{2k!} (i\boldsymbol{\lambda})^{\otimes 2k} \right) \sin\left(\sum_{k=2}^{\infty} \frac{\mathbf{c}(2k-1,d)'}{(2k-1)!} (-i)^{2k} (\boldsymbol{\lambda})^{\otimes(2k-1)} \right) = \\ & \frac{\mathbf{c}(3,d)'\boldsymbol{\lambda}^{\otimes 3}}{3!} - \frac{\mathbf{c}(5,d)'\boldsymbol{\lambda}^{\otimes 5}}{5!} + \dots \end{aligned} \tag{41}$$

Using above equations (40) and (41), the equation (39) can be re-written as:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\left\{\exp\left(-\frac{\mathbf{c}(2,d)'}{2}\boldsymbol{\lambda}^{\otimes 2}\right)\cos((\mathbf{x}-\mathbf{c}(1,d))'\boldsymbol{\lambda})\right.$$
$$\left.\left(1+\frac{\mathbf{c}(4,d)'\boldsymbol{\lambda}^{\otimes 4}}{4!}-\frac{\mathbf{c}(6,d)'\boldsymbol{\lambda}^{\otimes 6}}{6!}-\frac{1}{2}\left(\frac{\mathbf{c}(3,d)'\boldsymbol{\lambda}^{\otimes 3}}{3!}\right)^{\otimes 2}+\ldots\right)\right\}d\boldsymbol{\lambda}$$
$$-\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\left\{\exp\left(-\frac{\mathbf{c}(2,d)'}{2}\boldsymbol{\lambda}^{\otimes 2}\right)\cos\left((\mathbf{x}-\mathbf{c}(1,d))'\boldsymbol{\lambda}-\frac{\pi}{2}\right)\right.$$
$$\left.\left(\frac{\mathbf{c}(3,d)'\boldsymbol{\lambda}^{\otimes 3}}{3!}-\frac{\mathbf{c}(5,d)'\boldsymbol{\lambda}^{\otimes 5}}{5!}+\ldots\right)\right\}d\boldsymbol{\lambda} \tag{42}$$

Using the equation (30) for derivatives of Gaussian defined, the above equation can be simplified as:

$$f(\mathbf{x}) = G(\mathbf{x}) - \frac{\mathbf{c}(3,d)'}{3!}G^{(3)}(\mathbf{x}) + \frac{\mathbf{c}(4,d)'}{4!}G^{(4)}(\mathbf{x}) - \frac{\mathbf{c}(5,d)'}{5!}G^{(5)}(\mathbf{x})$$
$$+\frac{\mathbf{c}(6,d)'+10\mathbf{c}(3,d)^{\otimes 2'}}{6!}G^{(6)}(\mathbf{x})+\ldots \tag{43}$$

The equation (43) is the Gram-Charlier A series expressed directly in terms of the cumulants and the derivatives of the Gaussian PDF. Usually, the GCA is represented in terms of the Hermite polynomials. So, the GCA expansion in terms of the Hermite polynomials, derived in equation (34), can be re-written as:

$$f(\mathbf{x}) = G(\mathbf{x})\left[1+\frac{\mathbf{c}(3,d)'}{3!}\left(\mathbf{C}_{\mathbf{x}}^{-1}\right)^{\otimes 3}\mathbf{H}_3(\mathbf{x};\mathbf{0},\mathbf{C}_{\mathbf{x}}) + \frac{\mathbf{c}(4,d)'}{4!}\left(\mathbf{C}_{\mathbf{x}}^{-1}\right)^{\otimes 4}\mathbf{H}_4(\mathbf{x};\mathbf{0},\mathbf{C}_{\mathbf{x}})\right.$$
$$\left.+\frac{\mathbf{c}(5,d)'}{5!}\left(\mathbf{C}_{\mathbf{x}}^{-1}\right)^{\otimes 5}\mathbf{H}_5(\mathbf{x};\mathbf{0},\mathbf{C}_{\mathbf{x}}) + \frac{\mathbf{c}(6,d)'+10\mathbf{c}(3,d)^{\otimes 2'}}{6!}\left(\mathbf{C}_{\mathbf{x}}^{-1}\right)^{\otimes 6}\mathbf{H}_6(\mathbf{x};\mathbf{0},\mathbf{C}_{\mathbf{x}})+\ldots\right] \tag{44}$$

$$f(\mathbf{x}) = G(\mathbf{x})\left[1+\frac{\mathbf{c}(3,d)'}{3!}\mathbf{H}_3(\mathbf{x};\mathbf{0},\mathbf{I}_d) + \frac{\mathbf{c}(4,d)'}{4!}\mathbf{H}_4(\mathbf{x};\mathbf{0},\mathbf{I}_d) + \frac{\mathbf{c}(5,d)'}{5!}\mathbf{H}_5(\mathbf{x};\mathbf{0},\mathbf{I}_d)\right.$$
$$\left.+\frac{\mathbf{c}(6,d)'+10\mathbf{c}(3,d)^{\otimes 2'}}{6!}\mathbf{H}_6(\mathbf{x};\mathbf{0},\mathbf{I}_d)+\ldots\right] \tag{45}$$

7. Multivariate Generalized Gram-Charlier series

To derive the generalized Gram-Charlier series, an unknown PDF $f(\mathbf{x})$ need be represented in terms of any known reference PDF $\psi(\mathbf{x})$, where both the PDFs are represented in terms of their cumulants. Let the k^{th} order cumulant vector of the reference PDF $\psi(\mathbf{x})$ be $\mathbf{c}_r(k,d)$. Then, the k^{th} order cumulant difference vector $\boldsymbol{\delta}(k,d)$ is given by $\boldsymbol{\delta}(k,d) = \mathbf{c}(k,d) - \mathbf{c}_r(k,d), \forall k$. Using this cumulant difference vector, the equation (25) can be re-

written as under:

$$f(\mathbf{x}) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \exp\left(\sum_{k=1}^{\infty} \frac{\mathbf{c}_r(2k,d)'}{2k!}(i\boldsymbol{\lambda})^{\otimes 2k}\right) \exp\left(\sum_{k=1}^{\infty} \frac{\boldsymbol{\delta}(2k,d)'}{2k!}(i\boldsymbol{\lambda})^{\otimes 2k}\right)$$
$$\cos\left(\mathbf{x}'\boldsymbol{\lambda} + \sum_{k=1}^{\infty} \frac{\mathbf{c}_r(2k-1,d)'}{(2k-1)!}(i)^{2k}(\boldsymbol{\lambda})^{\otimes(2k-1)} + \sum_{k=1}^{\infty} \frac{\boldsymbol{\delta}(2k-1,d)'}{(2k-1)!}(i)^{2k}(\boldsymbol{\lambda})^{\otimes(2k-1)}\right) d\boldsymbol{\lambda} \tag{46}$$

$$= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \exp\left(\sum_{k=1}^{\infty} \frac{\mathbf{c}_r(2k,d)'}{2k!}(i\boldsymbol{\lambda})^{\otimes 2k}\right) \exp\left(\sum_{k=1}^{\infty} \frac{\boldsymbol{\delta}(2k,d)'}{2k!}(i\boldsymbol{\lambda})^{\otimes 2k}\right)$$
$$\left\{\cos\left(\mathbf{x}'\boldsymbol{\lambda} + \sum_{k=1}^{\infty} \frac{\mathbf{c}_r(2k-1,d)'}{(2k-1)!}(i)^{2k}(\boldsymbol{\lambda})^{\otimes(2k-1)}\right) \cos\left(\sum_{k=1}^{\infty} \frac{\boldsymbol{\delta}(2k-1,d)'}{(2k-1)!}(i)^{2k}(\boldsymbol{\lambda})^{\otimes(2k-1)}\right)\right.$$
$$\left. - \sin\left(\mathbf{x}'\boldsymbol{\lambda} + \sum_{k=1}^{\infty} \frac{\mathbf{c}_r(2k-1,d)'}{(2k-1)!}(i)^{2k}(\boldsymbol{\lambda})^{\otimes(2k-1)}\right) \sin\left(\sum_{k=1}^{\infty} \frac{\boldsymbol{\delta}(2k-1,d)'}{(2k-1)!}(i)^{2k}(\boldsymbol{\lambda})^{\otimes(2k-1)}\right)\right\} d\boldsymbol{\lambda} \tag{47}$$

Using the expansions in Appendix B, parts of the equation (47) can be approximated upto maximum 6^{th}-order statistics as under:

$$
\begin{aligned}
&\exp\left(\sum_{k=1}^{\infty}\frac{\boldsymbol{\delta}(2k,d)'}{2k!}(i\boldsymbol{\lambda})^{\otimes 2k}\right)\cos\left(\sum_{k=1}^{\infty}\frac{\boldsymbol{\delta}(2k-1,d)'}{(2k-1)!}(i)^{2k}(\boldsymbol{\lambda})^{\otimes(2k-1)}\right)\\
&= 1+\left(\sum_{k=1}^{\infty}\frac{\boldsymbol{\delta}(2k,d)'}{2k!}(i\boldsymbol{\lambda})^{\otimes 2k}\right)+\frac{1}{2}\left(\sum_{k=1}^{\infty}\frac{\boldsymbol{\delta}(2k,d)'}{2k!}(i\boldsymbol{\lambda})^{\otimes 2k}\right)^{\otimes 2}\\
&\quad-\frac{1}{2}\left(\sum_{k=1}^{\infty}\frac{\boldsymbol{\delta}(2k-1,d)'}{(2k-1)!}(i)^{2k}(\boldsymbol{\lambda})^{\otimes(2k-1)}\right)^{\otimes 2}-\ldots\\
&= 1+\left(-\frac{\boldsymbol{\delta}(2,d)'}{2!}\boldsymbol{\lambda}^{\otimes 2}+\frac{\boldsymbol{\delta}(4,d)'}{4!}\boldsymbol{\lambda}^{\otimes 4}-\frac{\boldsymbol{\delta}(6,d)'}{6!}\boldsymbol{\lambda}^{\otimes 6}\right)+\frac{1}{2}\left(-\frac{\boldsymbol{\delta}(2,d)'}{2!}\boldsymbol{\lambda}^{\otimes 2}+\frac{\boldsymbol{\delta}(4,d)'}{4!}\boldsymbol{\lambda}^{\otimes 4}\right)^{\otimes 2}\\
&\quad-\frac{1}{2}\left(-\boldsymbol{\delta}(1,d)'\boldsymbol{\lambda}+\frac{\boldsymbol{\delta}(3,d)'}{3!}\boldsymbol{\lambda}^{\otimes 3}-\frac{\boldsymbol{\delta}(5,d)'}{5!}\boldsymbol{\lambda}^{\otimes 5}\right)^{\otimes 2}+\frac{1}{6}\left(-\frac{\boldsymbol{\delta}(2,d)'}{2!}\boldsymbol{\lambda}^{\otimes 2}\right)^{\otimes 3}\\
&\quad-\frac{1}{2}\left(\left(-\frac{\boldsymbol{\delta}(2,d)'}{2!}\boldsymbol{\lambda}^{\otimes 2}+\frac{\boldsymbol{\delta}(4,d)'}{4!}\boldsymbol{\lambda}^{\otimes 4}\right)\otimes\left(\boldsymbol{\delta}(1,d)'\boldsymbol{\lambda}\right)^{\otimes 2}\right)\\
&\quad-\frac{1}{4}\left(\left(-\frac{\boldsymbol{\delta}(2,d)'}{2!}\boldsymbol{\lambda}^{\otimes 2}\right)^{\otimes 2}\otimes\left(\boldsymbol{\delta}(1,d)'\boldsymbol{\lambda}\right)^{\otimes 2}\right)+\frac{1}{4!}\left(-\frac{\boldsymbol{\delta}(2,d)'}{2!}\boldsymbol{\lambda}^{\otimes 2}\right)^{\otimes 4}\\
&\quad-\frac{1}{4!}\left(-\boldsymbol{\delta}(1,d)'\boldsymbol{\lambda}\right)^{\otimes 4}-\left(\frac{\boldsymbol{\delta}(2,d)'}{2!}\boldsymbol{\lambda}^{\otimes 2}\otimes\boldsymbol{\delta}(1,d)'\boldsymbol{\lambda}^{\otimes 4}\right)+\ldots\\
&= 1-\frac{1}{2}\left(\boldsymbol{\delta}(1,d)^{\otimes 2}+\boldsymbol{\delta}(2,d)\right)'\boldsymbol{\lambda}^{\otimes 2}\\
&\quad+\frac{1}{4!}\left(\boldsymbol{\delta}(1,d)^{\otimes 4}+6\boldsymbol{\delta}(2,d)\otimes\boldsymbol{\delta}(1,d)^{\otimes 2}+3\boldsymbol{\delta}(2,d)^{\otimes 2}+4\boldsymbol{\delta}(3,d)\otimes\boldsymbol{\delta}(1,d)+\boldsymbol{\delta}(4,d)\right)'\boldsymbol{\lambda}^{\otimes 4}\\
&\quad-\frac{1}{6!}\left(\boldsymbol{\delta}(1,d)^{\otimes 6}+15\boldsymbol{\delta}(2,d)^{\otimes 3}+10\boldsymbol{\delta}(3,d)^{\otimes 2}+15\boldsymbol{\delta}(4,d)\otimes\boldsymbol{\delta}(2,d)+15\boldsymbol{\delta}(4,d)\otimes\boldsymbol{\delta}(1,d)^{\otimes 2}\right.\\
&\quad+20\boldsymbol{\delta}(3,d)\otimes\boldsymbol{\delta}(1,d)^{\otimes 2}+15\boldsymbol{\delta}(2,d)\otimes\boldsymbol{\delta}(1,d)^{\otimes 4}+45\boldsymbol{\delta}(2,d)^{\otimes 2}\otimes\boldsymbol{\delta}(1,d)^{\otimes 2}\\
&\quad\left.+6\boldsymbol{\delta}(5,d)\otimes\boldsymbol{\delta}(1,d)+60\boldsymbol{\delta}(3,d)\otimes\boldsymbol{\delta}(2,d)\otimes\boldsymbol{\delta}(1,d)+\boldsymbol{\delta}(6,d)\right)'\boldsymbol{\lambda}^{\otimes 6}+\ldots
\end{aligned}
\tag{48}
$$

Similarly,

$$\begin{aligned}
&\exp\left(\sum_{k=1}^{\infty}\frac{\boldsymbol{\delta}(2k,d)'}{2k!}(i\boldsymbol{\lambda})^{\otimes 2k}\right)\sin\left(\sum_{k=1}^{\infty}\frac{\boldsymbol{\delta}(2k-1,d)'}{(2k-1)!}(-i)^{2k}(\boldsymbol{\lambda})^{\otimes(2k-1)}\right)\\
=\;&\left(-\boldsymbol{\delta}(1,d)'\boldsymbol{\lambda}+\frac{\boldsymbol{\delta}(3,d)'}{3!}\boldsymbol{\lambda}^{\otimes 3}-\frac{\boldsymbol{\delta}(5,d)'}{5!}\boldsymbol{\lambda}^{\otimes 5}\right)+\left(\frac{\boldsymbol{\delta}(2,d)'\boldsymbol{\lambda}^{\otimes 2}}{2!}\otimes\frac{\boldsymbol{\delta}(1,d)'\boldsymbol{\lambda}}{1}\right.\\
&\left.+\frac{\boldsymbol{\delta}(3,d)'\boldsymbol{\lambda}^{\otimes 3}}{3!}\otimes\frac{\boldsymbol{\delta}(2,d)'\boldsymbol{\lambda}^{\otimes 2}}{2!}+\frac{\boldsymbol{\delta}(4,d)'\boldsymbol{\lambda}^{\otimes 4}}{4!}\otimes\frac{\boldsymbol{\delta}(1,d)'\boldsymbol{\lambda}^{\otimes 1}}{1!}\right)\\
&-\frac{1}{2}\left(\frac{\boldsymbol{\delta}(2,d)'\boldsymbol{\lambda}^{\otimes 2}}{2!}\right)^{\otimes 2}\otimes\frac{\boldsymbol{\delta}(1,d)'\boldsymbol{\lambda}^{\otimes 1}}{1!}-\left(-\boldsymbol{\delta}(1,d)'\boldsymbol{\lambda}\right)^{\otimes 3}+\ldots\\
=\;&\boldsymbol{\delta}(1,d)'\boldsymbol{\lambda}+\frac{1}{3!}\left(\boldsymbol{\delta}(3,d)+3\boldsymbol{\delta}(2,d)\otimes\boldsymbol{\delta}(1,d)+\boldsymbol{\delta}(1,d)^{\otimes 3}\right)'\boldsymbol{\lambda}^{\otimes 3}\\
&-\frac{1}{5!}\left(\boldsymbol{\delta}(5,d)+5\boldsymbol{\delta}(4,d)\otimes\boldsymbol{\delta}(1,d)+15\boldsymbol{\delta}(2,d)^{\otimes 2}\otimes\boldsymbol{\delta}(1,d)+10\boldsymbol{\delta}(3,d)\otimes\boldsymbol{\delta}(2,d)\right.\\
&\left.+10\boldsymbol{\delta}(2,d)\otimes\boldsymbol{\delta}(1,d)^{\otimes 3}+10\boldsymbol{\delta}(3,d)\otimes\boldsymbol{\delta}(1,d)^{\otimes 2}+\boldsymbol{\delta}(1,d)^{\otimes 5}\right)'\boldsymbol{\lambda}^{\otimes 5}\ldots
\end{aligned}\tag{49}$$

Similar to equation (30) for the derivative of the Gaussian, the k^{th} derivative of the reference PDF can be obtained by taking k^{th}-order K-derivative of equation (25), as under:

$$\begin{aligned}
\psi^{(k)}(\mathbf{x})=\mathbf{D}_{\mathbf{x}}^{\otimes k}\psi(\mathbf{x})=\frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}(\boldsymbol{\lambda})^{\otimes k}\exp\left(\sum_{k=1}^{\infty}\frac{\mathbf{c}_r(2k,d)'}{2k!}(i\boldsymbol{\lambda})^{\otimes 2k}\right)\\
\cos\left(\mathbf{x}'\boldsymbol{\lambda}+\sum_{k=1}^{\infty}\frac{\mathbf{c}_r(2k-1,d)'}{(2k-1)!}(i)^{2k}(\boldsymbol{\lambda})^{\otimes(2k-1)}+\frac{k\pi}{2}\right)d\boldsymbol{\lambda}
\end{aligned}\tag{50}$$

The above equation (50) with the previous results on expansions in equation (48) and (49) can be used to simplify the equation (47). This derives the Generalized Gram-Charlier (GGC) series in terms of the cumulant difference ($\boldsymbol{\delta}(k,d)$) as under:

$$f(\mathbf{x})=\sum_{k=0}^{\infty}(-1)^k\frac{\boldsymbol{\alpha}(k,d)'}{k!}\psi^{(k)}(\mathbf{x})\tag{51}$$

where,

$$\begin{aligned}
\boldsymbol{\alpha}(0,d)=&1\\
\boldsymbol{\alpha}(1,d)=&\boldsymbol{\delta}(1,d)\\
\boldsymbol{\alpha}(2,d)=&\boldsymbol{\delta}(2,d)+\boldsymbol{\delta}(1,d)^{\otimes 2}\\
\boldsymbol{\alpha}(3,d)=&\boldsymbol{\delta}(3,d)+3\boldsymbol{\delta}(2,d)\otimes\boldsymbol{\delta}(1,d)+\boldsymbol{\delta}(1,d)^{\otimes 3}\\
\boldsymbol{\alpha}(4,d)=&\boldsymbol{\delta}(4,d)+4\boldsymbol{\delta}(3,d)\otimes\boldsymbol{\delta}(1,d)+3\boldsymbol{\delta}(2,d)^{\otimes 2}+6\boldsymbol{\delta}(2,d)\otimes\boldsymbol{\delta}(1,d)^{\otimes 2}+\boldsymbol{\delta}(1,d)^{\otimes 4}\\
\boldsymbol{\alpha}(5,d)=&\boldsymbol{\delta}(5,d)+5\boldsymbol{\delta}(4,d)\otimes\boldsymbol{\delta}(1,d)+10\boldsymbol{\delta}(3,d)\otimes\boldsymbol{\delta}(2,d)+10\boldsymbol{\delta}(3,d)\otimes\boldsymbol{\delta}(1,d)^{\otimes 2}\\
&+15\boldsymbol{\delta}(2,d)^{\otimes 2}\otimes\boldsymbol{\delta}(1,d)+10\boldsymbol{\delta}(2,d)\otimes\boldsymbol{\delta}(1,d)^{\otimes 3}+\boldsymbol{\delta}(1,d)^{\otimes 5}\\
\boldsymbol{\alpha}(6,d)=&\boldsymbol{\delta}(6,d)+6\boldsymbol{\delta}(5,d)\otimes\boldsymbol{\delta}(1,d)+15\boldsymbol{\delta}(4,d)\otimes\boldsymbol{\delta}(2,d)+15\boldsymbol{\delta}(4,d)\otimes\boldsymbol{\delta}(1,d)^{\otimes 2}+10\boldsymbol{\delta}(3,d)^{\otimes 2}\\
&+60\boldsymbol{\delta}(3,d)\otimes\boldsymbol{\delta}(2,d)\otimes\boldsymbol{\delta}(1,d)+20\boldsymbol{\delta}(3,d)\otimes\boldsymbol{\delta}(1,d)^{\otimes 2}+15\boldsymbol{\delta}(2,d)^{\otimes 3}\\
&+45\boldsymbol{\delta}(2,d)^{\otimes 2}\otimes\boldsymbol{\delta}(1,d)^{\otimes 2}+15\boldsymbol{\delta}(2,d)\otimes\boldsymbol{\delta}(1,d)^{\otimes 4}+\boldsymbol{\delta}(1,d)^{\otimes 6}
\end{aligned}\tag{52}$$

The above equations have exact resemblance with those expressing moments in terms of the cumulants in Section 3.1. This must happen, as equation (51) for GGC expansion is matching that for the PDF representation in terms of the derivative of delta function in equation (12). This matching proves that $\boldsymbol{\alpha}(k,d)$ is related in same way to $\boldsymbol{\delta}(k,d)$, as $\mathbf{m}(k,d)$ to $\mathbf{c}(k,d)$.

The verification of the derived GGC can be obtained by taking Gaussian density as a reference PDF. With Gaussian density as a reference, $\boldsymbol{\delta}(1,d) = \mathbf{0}$ and $\boldsymbol{\delta}(2,d) = \mathbf{0}$. So, the coefficients $\boldsymbol{\alpha}(k,d)$ in equation (51) can be derived as under:

$$\begin{aligned}
\boldsymbol{\alpha}(0,d) =& 1 \\
\boldsymbol{\alpha}(1,d) =& \mathbf{0} \\
\boldsymbol{\alpha}(2,d) =& \mathbf{0} \\
\boldsymbol{\alpha}(3,d) =& \boldsymbol{\delta}(3,d) = \mathbf{c}(3,d) \\
\boldsymbol{\alpha}(4,d) =& \boldsymbol{\delta}(4,d) = \mathbf{c}(4,d) \\
\boldsymbol{\alpha}(5,d) =& \boldsymbol{\delta}(5,d) = \mathbf{c}(5,d) \\
\boldsymbol{\alpha}(6,d) =& \boldsymbol{\delta}(6,d) + 10\boldsymbol{\delta}(3,d)^{\otimes 2} = \mathbf{c}(6,d) + 10\mathbf{c}(3,d)^{\otimes 2}
\end{aligned} \tag{53}$$

Thus, the GGC series is derived and verified using known examples.

8. Characteristic function of an unknown random vector in terms of a reference characteristic function

The GGC derived as in equation (51) can be used to give the characteristic function of an unknown PDF, in terms of the characteristic function of a reference PDF. For that taking Fourier transform of equation (51), we get:

$$\mathcal{F}_{\mathbf{x}}(\boldsymbol{\lambda}) = \sum_{k=0}^{\infty} (-1)^k \frac{\boldsymbol{\alpha}(k,d)'}{k!} \mathsf{F}(\psi^{(k)}(\mathbf{x})) \tag{54}$$

Using differentiation property of Fourier transform,

$$\mathcal{F}_{\mathbf{x}}(\boldsymbol{\lambda}) = \left[\sum_{k=0}^{\infty} \boldsymbol{\alpha}(k,d)' \frac{(i\boldsymbol{\lambda})^{\otimes k}}{k!}\right] \mathsf{F}(\psi(\mathbf{x})) \tag{55}$$

$$= \exp\left[\sum_{k=1}^{\infty} \boldsymbol{\delta}(k,d)' \frac{(i\boldsymbol{\lambda})^{\otimes k}}{k!}\right] \mathcal{F}_r(\mathbf{x}) \tag{56}$$

where, $\mathcal{F}_r$ is the characteristic function of the reference PDF.

9. Compact derivation for the Generalized Gram-Charlier expansion

The compact derivation of equation (51) follows as under:

$$\mathcal{F}_{\mathbf{x}}(\boldsymbol{\lambda}) = \exp\left[\sum_{k=1}^{\infty} \mathbf{c}(k,d)' \frac{(i\boldsymbol{\lambda})^{\otimes k}}{k!}\right] \quad (\because \text{definition in equation (17)}) \tag{57}$$

$$= \exp\left[\sum_{k=1}^{\infty} \boldsymbol{\delta}(k,d)' \frac{(i\boldsymbol{\lambda})^{\otimes k}}{k!}\right] \exp\left[\sum_{k=1}^{\infty} \mathbf{c}_r(k,d) \frac{(i\boldsymbol{\lambda})^{\otimes k}}{k!}\right] \quad (\because \boldsymbol{\delta}(k,d) = \mathbf{c}(k,d) - \mathbf{c}_r(k,d)) \tag{58}$$

$$= \left[\sum_{k=0}^{\infty} \boldsymbol{\alpha}(k,d)' \frac{(i\boldsymbol{\lambda})^{\otimes k}}{k!}\right] \mathsf{F}(\psi(\mathbf{x})) \tag{59}$$

Taking inverse Fourier transform of the above equation brings

$$f(\mathbf{x}) = \sum_{k=0}^{\infty} \boldsymbol{\alpha}(k,d)' \frac{(-1)^k}{k!} \boldsymbol{\delta}^{(k)}(\mathbf{x}) * \psi(\mathbf{x}) \tag{60}$$

$$\text{or } f(\mathbf{x}) = \sum_{k=0}^{\infty} \boldsymbol{\alpha}(k,d)' \frac{(-1)^k}{k!} \psi^{(k)}(\mathbf{x}) \tag{61}$$

where, $*$ indicates convolution. Thus, the equation (51) is obtained in a more compact way.

10. Conclusion

The article has derived multivariate Generalized Gram-Charlier (GGC) expansion in equation (51) representing an unknown multivariate PDF in terms of the differences between cumulant vectors of an unknown PDF and cumulant vectors of the known reference PDF. The multivariate Gram-Charlier A series is derived in equation (43) representing an unknown multivariate PDF in terms of its cumulants. There has been also obtained multivariate moment and cumulant relationship in equation (21); the integral form of multivariate PDF representation in equation (25) and the integral form of multivariate Hermite polynomials in equation (34), equation (35). The expressions are derived using only elementary calculus of several variables through Kronecker product based derivative operator. so, the derived multivariate expressions are using vector notations and apparently similar to those for univariate.

Appendix A. Some important properties of the K-derivative operator

Some important properties of the K-derivative operator are listed below.

Property 1 (Scaling Property) *Let* $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_d)'$, $\boldsymbol{\lambda} \in \mathbb{R}^d$, $\mathbf{f}(\boldsymbol{\lambda}) \in \mathbb{R}^m$ *and* $\mathbf{f}_1(\boldsymbol{\lambda}) = \mathbf{A}\mathbf{f}(\boldsymbol{\lambda})$, *where* $\mathbf{A}$ *is an* $n \times m$ *matrix. Then*

$$\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes}(\mathbf{f}_1) = (\mathbf{A} \otimes \mathbf{I}_d)\, \mathbf{D}_{\boldsymbol{\lambda}}^{\otimes}(\mathbf{f}) \tag{62}$$

where, $\mathbf{I}_d$ *is a d-dimensional unit matrix.*

Property 2 (Chain Rule) *Let* $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_d)'$, $\boldsymbol{\lambda} \in \mathbb{R}^d$, $\mathbf{f}(\boldsymbol{\lambda}) \in \mathbb{R}^{m_1}$ *and* $\mathbf{g}(\boldsymbol{\lambda}) \in \mathbb{R}^{m_2}$. *Then*

$$\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes}(\mathbf{f} \otimes \mathbf{g}) = \mathbf{K}_{3\leftrightarrow 2}^{-1}(m_1, m_2, d)((\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes}\mathbf{f}) \otimes \mathbf{g}) + \mathbf{f} \otimes (\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes}\mathbf{g}) \tag{63}$$

where, $\mathbf{K}_{3\leftrightarrow 2}(m_1, m_2, d)$ *is a commutation matrix of size* $m_1 m_2 d \times m_1 m_2 d$ *that changes the order of the the Kronecker product components. For example,*

$$\mathbf{K}_{3\leftrightarrow 2}(m_1, m_2, d)(\mathbf{a_1} \otimes \mathbf{a_2} \otimes \mathbf{a_3}) = \mathbf{a_1} \otimes \mathbf{a_3} \otimes \mathbf{a_2}$$

The above K-derivative properties can be used to derive further the following properties:

1. $\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes}\boldsymbol{\lambda} = Vec\ \mathbf{I}_d$

2. $\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes}\boldsymbol{\lambda}^{\otimes k} = k\left(\boldsymbol{\lambda}^{\otimes k-1} \otimes Vec\ \mathbf{I}_d\right)$

3. Let $g(\boldsymbol{\lambda}) = \mathbf{a}' f(\boldsymbol{\lambda})$
   $\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes} g(\boldsymbol{\lambda}) = (\mathbf{a} \otimes \mathbf{I}_d)'\, \mathbf{D}_{\boldsymbol{\lambda}}^{\otimes} f(\boldsymbol{\lambda})$

4. $\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes}\mathbf{a}'\boldsymbol{\lambda}^{\otimes k} = k\,(\mathbf{a} \otimes \mathbf{I}_d)'\left(\boldsymbol{\lambda}^{\otimes k-1} \otimes Vec\ \mathbf{I}_d\right) = k\ \mathbf{a}'\left(\boldsymbol{\lambda}^{\otimes k-1} \otimes \mathbf{I}_d\right)$

5. $\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes}\mathbf{a}'^{\otimes k}\boldsymbol{\lambda}^{\otimes k} = k\left(\mathbf{a}^{\otimes k} \otimes \mathbf{I}_d\right)'\left(\boldsymbol{\lambda}^{\otimes k-1} \otimes Vec\ \mathbf{I}_d\right) = k\,(\mathbf{a}'\boldsymbol{\lambda})^{\otimes k-1} \otimes \mathbf{a}$

6. $\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes r}\boldsymbol{\lambda}^{\otimes k} = k(k-1)\cdots(k-r+1)\left(\boldsymbol{\lambda}^{\otimes k-r} \otimes (Vec\ \mathbf{I}_d)^{\otimes r}\right)$
   ($\because$ Repeated application of the Chain Rule)

7. $\mathbf{D}_{\boldsymbol{\lambda}}^{\otimes r}\mathbf{a}'^{\otimes k}\boldsymbol{\lambda}^{\otimes k} = k(k-1)\cdots(k-r+1)\left(\mathbf{a}^{\otimes k} \otimes \mathbf{I}_d^{\otimes r}\right)'\left(\boldsymbol{\lambda}^{\otimes k-r} \otimes (Vec\ \mathbf{I}_d)^{\otimes r}\right)$
   $= k(k-1)\cdots(k-r+1)\,(\mathbf{a}'\boldsymbol{\lambda})^{\otimes k-r} \otimes \mathbf{a}^{\otimes r}$
   ($\because$ Repeated application of the Chain Rule and the property:)

Appendix B. The Taylor series expansion of some required functions near zero

The Taylor series expansion of the required functions near $\boldsymbol{\lambda} = \mathbf{0}$, based on the equation (7), are given as under:

$$\mathrm{e}^{\mathbf{a}'\mathbf{x}} = \sum_{k=0}^{\infty} \frac{(\mathbf{a}'\mathbf{x})^{\otimes k}}{k!} = 1 + \mathbf{a}'\mathbf{x} + \frac{(\mathbf{a}'\mathbf{x})^{\otimes 2}}{2!} + \frac{(\mathbf{a}'\mathbf{x})^{\otimes 3}}{3!} + \dots \tag{64}$$

$$\sin(\mathbf{a}'\mathbf{x}) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}(\mathbf{a}'\mathbf{x})^{\otimes 2k+1} = \mathbf{a}'\mathbf{x} - \frac{(\mathbf{a}'\mathbf{x})^{\otimes 3}}{3!} + \frac{(\mathbf{a}'\mathbf{x})^{\otimes 5}}{5!} - \dots \tag{65}$$

$$\cos(\mathbf{a}'\mathbf{x}) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}(\mathbf{a}'\mathbf{x})^{\otimes 2k} = 1 - \frac{(\mathbf{a}'\mathbf{x})^{\otimes 2}}{2!} + \frac{(\mathbf{a}'\mathbf{x})^{\otimes 4}}{4!} - \dots \tag{66}$$

$$\begin{aligned}\mathrm{e}^{\mathbf{a}'\mathbf{x}}\cos(\mathbf{b}'\mathbf{y}) = 1 + \mathbf{a}'\mathbf{x} + \frac{(\mathbf{a}'\mathbf{x})^{\otimes 2}}{2!} - \frac{(\mathbf{b}'\mathbf{y})^{\otimes 2}}{2!} + \frac{(\mathbf{a}'\mathbf{x})^{\otimes 3}}{3!} - \frac{(\mathbf{a}'\mathbf{x}) \otimes (\mathbf{b}'\mathbf{y})^{\otimes 2}}{2!} \\ + \frac{(\mathbf{a}'\mathbf{x})^{\otimes 4}}{4!} + \frac{(\mathbf{b}'\mathbf{y})^{\otimes 4}}{4!} - \frac{(\mathbf{a}'\mathbf{x})^{\otimes 2} \otimes (\mathbf{b}'\mathbf{y})^{\otimes 2}}{2!2!} + \dots\end{aligned} \tag{67}$$

$$\begin{aligned}\mathrm{e}^{\mathbf{a}'\mathbf{x}}\sin(\mathbf{b}'\mathbf{y}) = \mathbf{b}'\mathbf{y} + (\mathbf{a}'\mathbf{x}) \otimes (\mathbf{b}'\mathbf{y}) - \frac{(\mathbf{a}'\mathbf{x})^{\otimes 2} \otimes (\mathbf{b}'\mathbf{y})}{2!} - \frac{(\mathbf{b}'\mathbf{y})^{\otimes 3}}{3!} \\ - \frac{(\mathbf{a}'\mathbf{x}) \otimes (\mathbf{b}'\mathbf{y})^{\otimes 3}}{3!} + \frac{(\mathbf{b}'\mathbf{y})^{\otimes 5}}{5!} - \frac{(\mathbf{a}'\mathbf{x})^{\otimes 2} \otimes (\mathbf{b}'\mathbf{y})^{\otimes 3}}{2!3!} + \dots\end{aligned} \tag{68}$$

Appendix C. Relationship between $\mathbf{m}(k,d)$ and $\mathbf{c}(k,d)$ for $k = 2$ and $k = 3$

For $k = 2$, using equation (15) and equation (19) and the differentiation rules for Kronecker product the following is derived :

$$\mathbf{m}(2,d) = \mathbf{D}_{\boldsymbol{\lambda}}^{\otimes 2}\mathbf{M}(\boldsymbol{\lambda})\big|_{\boldsymbol{\lambda}=\mathbf{0}} \tag{69}$$

$$\begin{aligned}&\Rightarrow \sum_{p=2}^{\infty} \mathbf{m}(p,d)' \frac{\left(\boldsymbol{\lambda}^{\otimes(p-2)} \otimes \mathbf{I}_d \otimes \mathbf{I}_d\right)}{(p-2)!}\Bigg|_{\boldsymbol{\lambda}=\mathbf{0}} = \\ &\mathbf{K}^{-1}_{\mathfrak{p}3\leftrightarrow 2}(d_1,d_2,d)\left\{\left(\sum_{p=2}^{\infty} \mathbf{c}(p,d)' \frac{\left(\boldsymbol{\lambda}^{\otimes(p-2)} \otimes \mathbf{I}_d \otimes \mathbf{I}_d\right)}{(p-2)!}\right) \otimes \exp\left(\sum_{q=1}^{\infty} \mathbf{c}(q,d)' \frac{\boldsymbol{\lambda}^{\otimes q}}{q!}\right)\Bigg|_{\boldsymbol{\lambda}=\mathbf{0}}\right\} \\ &+ \left(\sum_{p=1}^{\infty} \mathbf{c}(p,d)' \frac{\left(\boldsymbol{\lambda}^{\otimes(p-1)} \otimes \mathbf{I}_d\right)}{(p-1)!}\right)^{\otimes 2} \otimes \exp\left(\sum_{q=1}^{\infty} \mathbf{c}(q,d)' \frac{\boldsymbol{\lambda}^{\otimes q}}{q!}\right)\Bigg|_{\boldsymbol{\lambda}=\mathbf{0}}\end{aligned}$$

$$\mathbf{m}(2,d) = \mathbf{c}(2,d) + \mathbf{c}(1,d)^{\otimes 2} \tag{70}$$

where, $\mathfrak{p}3 \leftrightarrow 2 \in \mathfrak{P}$ is the required permutation and $\mathbf{K}_{\mathfrak{p}3\leftrightarrow 2}(d_1,d_2,d)$ is the corresponding commutation matrix.
For $k = 3$, using equation (15) and equation (19) and the differentiation rules for Kronecker

product the following is derived:

$$\mathbf{m}(3,d) = \mathbf{D}_{\boldsymbol{\lambda}}^{\otimes 3}\mathbf{M}(\boldsymbol{\lambda})\big|_{\boldsymbol{\lambda}=\mathbf{0}} \tag{71}$$

$$\begin{aligned}
\Rightarrow \sum_{p=3}^{\infty}\mathbf{m}(p,d)'\frac{\left(\boldsymbol{\lambda}^{\otimes(p-3)}\otimes\mathbf{I}_{d^3}\right)}{(p-3)!}\Bigg|_{\boldsymbol{\lambda}=\mathbf{0}} &= \mathbf{K}_{\mathfrak{p}3\leftrightarrow 2}^{-1}(d_1,d_2,d)\mathbf{K}_{\mathfrak{p}3\leftrightarrow 2}^{-1}(d_1,d_2,d) \\
&\left(\sum_{p=3}^{\infty}\mathbf{c}(p,d)'\frac{\left(\boldsymbol{\lambda}^{\otimes(p-3)}\otimes\mathbf{I}_{d^3}\right)}{(p-3)!}\right)\otimes\exp\left(\sum_{q=1}^{\infty}\mathbf{c}(q,d)'\frac{\boldsymbol{\lambda}^{\otimes q}}{q!}\right)\Bigg|_{\boldsymbol{\lambda}=\mathbf{0}} \qquad (72)\\
&+\mathbf{K}_{\mathfrak{p}3\leftrightarrow 2}^{-1}(d_1,d_2,d)\left(\sum_{p=2}^{\infty}\mathbf{c}(p,d)'\frac{\left(\boldsymbol{\lambda}^{\otimes(p-2)}\otimes\mathbf{I}_{d^2}\right)}{(p-2)!}\right)\otimes\left(\sum_{p=1}^{\infty}\mathbf{c}(p,d)'\frac{\left(\boldsymbol{\lambda}^{\otimes(p-1)}\otimes\mathbf{I}_{d}\right)}{(p-1)!}\right)\\
&\otimes\exp\left(\sum_{q=1}^{\infty}\mathbf{c}(q,d)'\frac{\boldsymbol{\lambda}^{\otimes q}}{q!}\right)\Bigg|_{\boldsymbol{\lambda}=\mathbf{0}}+\mathbf{K}_{\mathfrak{p}3\leftrightarrow 2}^{-1}(d_1,d_2,d)2\left(\sum_{p=2}^{\infty}\mathbf{c}(p,d)'\frac{\left(\boldsymbol{\lambda}^{\otimes(p-2)}\otimes\mathbf{I}_{d^2}\right)}{(p-2)!}\right)^{\otimes 1}\\
&\otimes\left(\sum_{p=1}^{\infty}\mathbf{c}(p,d)'\frac{\left(\boldsymbol{\lambda}^{\otimes(p-1)}\otimes\mathbf{I}_{d}\right)}{(p-1)!}\right)^{\otimes 1}\\
&\otimes\exp\left(\sum_{q=1}^{\infty}\mathbf{c}(q,d)'\frac{\boldsymbol{\lambda}^{\otimes q}}{q!}\right)\Bigg|_{\boldsymbol{\lambda}=\mathbf{0}}+\left(\sum_{p=1}^{\infty}\mathbf{c}(p,d)'\frac{\left(\boldsymbol{\lambda}^{\otimes(p-1)}\otimes\mathbf{I}_{d}\right)}{(p-1)!}\right)^{\otimes 2}\\
&\otimes\left(\sum_{p=1}^{\infty}\mathbf{c}(p,d)'\frac{\left(\boldsymbol{\lambda}^{\otimes(p-1)}\otimes\mathbf{I}_{d}\right)}{(p-1)!}\right)^{\otimes 1}\otimes\exp\left(\sum_{q=1}^{\infty}\mathbf{c}(q,d)'\frac{\boldsymbol{\lambda}^{\otimes q}}{q!}\right)\Bigg|_{\boldsymbol{\lambda}=\mathbf{0}}
\end{aligned}$$

$$\mathbf{m}(3,d) = \mathbf{c}(3,d) + 3\mathbf{c}(2,d)\otimes\mathbf{c}(1,d) + \mathbf{c}(1,d)^{\otimes 3} \tag{73}$$

The above calculations signify the role of the commutation matrices in deciding the relations between moment vectors and cumulant vectors.

Appendix D. The proof related to Fourier Transform of multivariate Gaussian in Section 4.

Applying $\mathbf{D}_{\mathbf{x}}^{\otimes}$ on equation (28),

$$\mathbf{D}_{\mathbf{x}}^{\otimes}G(\mathbf{x}) = \frac{1}{(2\pi)^d}\int_{\mathbb{R}^d}\exp\left(-\frac{\mathbf{c}(2,d)'}{2}\boldsymbol{\lambda}^{\otimes 2}\right)\boldsymbol{\lambda}^{\otimes}\sin\left(\mathbf{x}'\boldsymbol{\lambda}-\mathbf{c}(1,d)'\boldsymbol{\lambda}\right)d\boldsymbol{\lambda} \tag{74}$$

$$\text{Now, } \mathbf{D}_{\boldsymbol{\lambda}}^{\otimes}\exp\left(-\frac{\mathbf{c}(2,d)'}{2}\boldsymbol{\lambda}^{\otimes 2}\right) = \mathbf{c}(2,d)'\left(\boldsymbol{\lambda}\otimes\mathbf{I}_d\right)\exp\left(-\frac{\mathbf{c}(2,d)'}{2}\boldsymbol{\lambda}^{\otimes 2}\right) \tag{75}$$

$$\text{Also, } \int_{\mathbb{R}^d}\cos\left(\mathbf{x}'\boldsymbol{\lambda}-\mathbf{c}(1,d)'\boldsymbol{\lambda}\right)d\boldsymbol{\lambda} = \frac{\sin\left(\mathbf{x}'\boldsymbol{\lambda}-\mathbf{c}(1,d)'\boldsymbol{\lambda}\right)}{\left(\mathbf{x}'\boldsymbol{\lambda}-\mathbf{c}(1,d)'\boldsymbol{\lambda}\right)} \tag{76}$$

Using above equations (75), (76) in equation (74) and applying integration by parts; we get:

$$\mathbf{D}_{\mathbf{x}}^{\otimes} G(\mathbf{x}) = -\left(inv\ \mathbf{c}(2,d)\right)' (\mathbf{x} - \mathbf{c}(1,d)) G(\mathbf{x}) \tag{77}$$

The solution of above differential equation leads to the following:

$$G(\mathbf{x}) = c \exp\left(-\left(inv\ \mathbf{c}(2,d)\right)' (\mathbf{x} - \mathbf{c}(1,d))\right) \tag{78}$$

$$\begin{aligned} \text{where, } c = G(\mathbf{0}) &= \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} \exp\left(-\frac{\mathbf{c}(2,d)'}{2} \boldsymbol{\lambda}^{\otimes 2}\right) d\boldsymbol{\lambda} \\ &= (2\pi)^{-d/2} \left(|\mathbf{C}_{\mathbf{x}}|\right)^{-1/2} \end{aligned} \tag{79}$$

$$\Rightarrow G(\mathbf{x}) = (2\pi)^{-d/2} \left(|\mathbf{C}_{\mathbf{x}}|\right)^{-1/2} \exp\left(-\left(inv\ \mathbf{c}(2,d)\right)' (\mathbf{x} - \mathbf{c}(1,d))\right) \tag{80}$$

References

Shun-ichi Amari and Masayuki Kumon. Differential geometry of edgeworth expansions in curved exponential family. *Annals Of The Institute Of Statistical Mathematics*, 35(1): 1–24, 1983.

Shun-ichi Amari, Andrzej Cichocki, Howard Hua Yang, et al. A new learning algorithm for blind signal separation. *Advances in neural information processing systems*, pages 757–763, 1996.

Leo A. Aroian. The type b gram-charlier series. *Ann. Math. Statist.*, 8(4):183–192, 12 1937.

Leo A. Aroian. The probability function of the product of two normally distributed variables. *The Annals of Mathematical Statistics*, 18(2):pp. 265–271, 1947.

Mario Berberan-Santos. Expressing a probability density function in terms of another pdf: A generalized gram-charlier expansion. *Journal of Mathematical Chemistry*, 42(3):585–594, October 2007.

S Berkowitz and FJ Garner. The calculation of multidimensional hermite polynomials and gram-charlier coefficients. *mathematics of computation*, pages 537–545, 1970.

R. Boscolo and V. P. Roychowdhury. On the uniqueness of the minimum of the information-theoretic cost function for the separation of mixtures of nearly gaussian signals. pages 137–141, Nara, Japan, April 2003.

Newton L Bowers. Expansion of probability density functions as a sum of gamma densities with applications in risk theory. *Trans. Soc. Actuaries*, 18(52):125, 1966.

Leon Cohen. Generalization of the gram-charlier/edgeworth series andapplication to time-frequency analysis. *Multidimensional Syst. Signal Process.*, 9(4):363–372, October 1998.

Leon Cohen. On the generalization of the edgeworth/gram-charlier series. *Journal of Mathematical Chemistry*, 49(3):625–628, March 2011.

A. W. Davis. Statistical distributions in univariate and multivariate edgeworth populations. *Biometrika*, 63(3):661–670, Dec., 1976.

Esther B Del Brio, Trino-Manuel Ñíguez, and Javier Perote. Gram-charlier densities: A multivariate approach. *Quantitative Finance*, 9(7):855–868, 2009.

Mark Girolami and Colin Fyfe. Negentropy and kurtosis as projection pursuit indices provide generalised ica algorithms. In *A. C, Back A (eds.), NIPS-96 Blind Signal Separation Workshop*, volume 8, 1996.

Anders Hald. The early history of the cumulants and the gram-charlier series. *International Statistical Review*, 68(2):137–153, 2000. ISSN 1751-5823.

Bjorn Holmquist. The d-variate vector hermite polynomial of order k. *Linear Algebra and its Applications*, 237-238(0):155–190, 1996. ISSN 0024-3795. Linear Algebra and Statistics: In Celebration of C. R. Rao's 75th Birthday (September 10, 1995).

A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. John Wiley, New York, 2001. 481+xxii pages.

Eric Jondeau and Michael Rockinger. Gram–charlier densities. *Journal of Economic Dynamics and Control*, 25(10):1457–1483, 2001.

Tonu Kollo and Dietrich von Rosen. A unified approach to the approximation of multivariate densities. *Scandinavian Journal of Statistics*, 25(1):pp. 93–109, 1998.

Jan R. Magnus. On the concept of matrix derivative. *Journal of multivariate analysis*, 101 (101):2200–2206, 2010.

Jan R. Magnus and Heinz Neudecker. *Matrix differential calculus with applications in statistics and econometrics*. J. Wiley & Sons, Chichester, New York, Weinheim, 1999.

P. McCullagh. *Tensor Methods in Statistics*. London; New York: Chapman & Hall, 1987.

HENRY LEWIS RIETZ. *Mathematical Statistics*, volume 3 of *Carus Mathematical Monographs*. Mathematical Association of America, 1 edition, 1927.

T. Subba Rao S. Rao Jammalamadaka and Gyorgy Terdik. Higher order cumulants of random vectors and applications to statistical inference and time series. *Sankhya: The Indian Journal of Statistics (2003-2007)*, 68(2):326–356, May, 2006.

P Sauer and G Heydt. A convenient multivariate gram-charlier type a series. *IEEE Transactions on Communications*, pages 247–248, 1979.

D. C. Schleher. Generalized gram-charlier series with application to the sum of log-normal variates (corresp.). *IEEE Transactions on Information Theory*, 23(2):275–280, 1977.

Ib M. Skovgaard. On multivariate edgeworth expansions. *International Statistical Review / Revue Internationale de Statistique*, 54(2):pp. 169–186, 1986.

Gyorgy Terdik. Higher order statistics and multivariate vector hermite polynomials. *Teor. Imovir. Mat. Stat.*, 66:147–168, 2002.

Larry A. Viehland. Velocity distribution functions and transport coefficients of atomic ions in atomic gases by a gramcharlier approach. *Chemical Physics*, 179(1):71 – 92, 1994.

Christopher S Withers and Saralees Nadarajah. The dual multivariate charlier and edgeworth expansions. *Statistics & Probability Letters*, 87:76–85, 2014.