Global Journal of Pure and Applied Mathematics.
ISSN 0973-1768 Volume 13, Number 11 (2017), pp. 8053–8064

http://www.ripublication.com/gjpam.htm

Multivariate Gaussian Estimation: The Principle of $n-$Cross Sections

A.C. Mkolesia[1]
Department of Mathematics and Statistics,
Tshwane University of Technology,
175 Nelson Mandela Drive, Pretoria, 0001, RSA.

C.R. Kikawa
Department of Mathematics and Statistics,
Namibia University of Science and Technology,
Windhoek, Namibia.

M.Y. Shatalov
Department of Mathematics and Statistics,
Tshwane University of Technology,
175 Nelson Mandela Drive, Pretoria, 0001, RSA.

Abstract

In this article, a new method, the principle of n-cross sections, for estimating the sufficient parameters of a multivariate Gaussian is proposed. The method computes, $\mathbf{\Sigma}$ and $\boldsymbol{\mu}$, by reducing the variable and parameter space $(\mathcal{H})$. The multivariate Gaussian is spliced into n-cross sections called hyperplanes $\left(\mathcal{P}_{ij}\right)$, at different angles (α_i), (in this work seven cross sections are used). The parameter space is reduced from six unknown parameters to three unknown parameters or from two unknowns to one unknown. Thus a minimum of three hyperplanes $\left(\mathcal{P}_{ij}\right)$ are required to compute the estimates for the multivariate Gaussian. The proposed method works optimally with a minimum of 50 data points for convergence. From the error analysis the proposed method (PnCS) performs better than the maximum likelihood estimator (MLE) method, with an error margin of 18.74% and 34.90% respectively. Thus PnCS is recommended for the estimation of the multivariate Gaussian distribution.

AMS subject classification: Primary 65C, 68U Secondary 62J.
Keywords: Multivariate Gaussian distributions, Parameter estimation, Hyperplanes estimation, Numerical Simulation.

[1]Corresponding author.

1. Introduction

Estimating parameters of a multivariate Gaussian distribution is usually done by methods maximum likelihood [1], [2], [3], which assumes that the training (available) data set is large enough to provide robust estimates. This study proposes a new method of parameter estimation for the multivariate Gaussian distribution known as the principle of n-cross sections (PCS). The multivariate Gaussian distribution is studied since several multivariate distributions rely in some way on the multivariate Gaussian [1], [2], [4]. It is stated that many real-world problems fall naturally within the framework of normal theory [1]. The importance of the normal distribution rests on its dual role as both population model for natural phenomena and approximate sampling distribution for many statistics [1], [2]. Real life data may never *exactly* come from the true multivariate Gaussian distribution [2], the multivariate Gaussian distribution provides a robust approximation to the "true"population distribution [2]. The multivariate distribution also has many "nice"mathematical properties that can be explored and mathematically tractable [2]. Ideally multivariate methods should be used so that the data can be treated in an integrated way [5]. Because of the central limiting theorem, many multivariate statistics converge to the multivariate Gaussian distribution as the sample sizes increase [2]. The multivariate distributions have a number of applications like, speaker authentication [1], decision theory [6], cluster analysis [6].

This paper is organized as follows, the importance of parameter estimation for a multivariate distribution is discussed in Section 1. In Section 2 the motivation for the study is discussed. Existing parameter estimation methods like, the maximum likelihood for multivariate Gaussian distributions is presented in Section 3.1. A theorem on which the PCS method is based is firstly stated and proved in Section 3.2. A numerical simulation is performed in Section 4 in order to evaluate the performance of the PCS method. In Section 5 a conclusion is presented.

2. Motivation

Much work on the estimation of multivariate Gaussian distributions has been done using nonparametric methods like the kernel estimators [7], [8]. Also there exists some traditional methods that are used to estimate the parameters for the multivariate Gaussian distribution like the maximum likelihood estimators [9], [10], [11], [12], however the MLE needs initialization [10], and has stringent distribution assumptions. There is a need to establish new methods that are more robust, have less distribution assumptions and don't require initialization which can estimate the parameters for multivariate Gaussian distributions. In this study a new method called the principle of n-cross section for estimating parameters of a multivariate Gaussian distribution is proposed.

Definition 2.1. Multivariate Gaussian: Any $\mathbb{R}^n$-valued random variable $\mathbf{x}$ is a multivariate Gaussian if for every $\mathbf{t} \in \mathbb{R}$ real valued random variable $\mathbf{t} \cdot \mathbf{x}$ is normal or Gaussian.

From definition (2.1) The distribution of the multivariate $\mathbf{t} \cdot \mathbf{x}$ is determined uniquely by its mean $\mu = \mu_{\mathbf{t}}$ and its standard deviation $\sigma = \sigma_{\mathbf{t}}$. Thus $\mu_{\mathbf{t}} = \mathbf{t} \cdot \boldsymbol{\mu}$ where $\boldsymbol{\mu} = \mathbb{E}(\mathbf{x})$ [2], [4].

A vector-valued random variable $\mathbf{x} = [\mathbf{x}_1, \ldots, \mathbf{x}_n]^T$ is said to be a multivariate normal (or Gaussian) distribution with mean (location parameter) $\boldsymbol{\mu} \in \mathbb{R}^n$ and covariance matrix $\boldsymbol{\Sigma} \in \mathbf{S}^n_{++}$, where

$$\mathbf{S}^n_{++} = \mathbf{A} \in \mathbb{R}^{n \times n} : \mathbf{A} = \mathbf{A}^T \text{ and } \mathbf{x}^T \mathbf{A}\, \mathbf{x} > 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n \text{ such that } \mathbf{x} \neq \mathbf{0},$$

that is to say $\mathbf{A}$ is a positive definite $n \times n$ matrix and the probability density function is given by

$$\mathcal{N}(\mathbf{x} \mid \boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}} \mid \Sigma \mid^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right] \quad \text{where} \quad \mathbf{x} \in \mathbb{R}^n. \tag{2.1}$$

Thus $\mathbf{x}$ is a Gaussian distribution, $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} \in \mathbb{R}^n$ and $\boldsymbol{\Sigma} > 0$ with moments, mean, $\mathbb{E}(\mathbf{x}) = \boldsymbol{\mu}$ and covariance, $\text{cov}(\mathbf{x}) = \boldsymbol{\Sigma}$ [13].

From Equation (2.1) with a mean vector $\boldsymbol{\mu}$ there are n (independent) parameters and within the symmetric covariance matrix $\boldsymbol{\Sigma}$ there are $\frac{1}{2}n(n+1)$ independent parameters, thus $\frac{1}{2}n(n+3)$ independent parameters in total [3].

3. Parameter Estimation for Multivariate Gaussian Distributions

There are several methods that are applied in order to estimate the parameters of multivariate Gaussian distributions. However in this section the MLE is considered, due to its robustness. The proposed PCS method is also presented.

3.1. Multivariate Gaussian in the MLE Framework.

Theorem 3.1. (MLE for a Gaussian) Suppose a set of n i.i.d. samples $\mathbf{x} = [\mathbf{x}_1, \ldots, \mathbf{x}_n]^T$ is sampled from a multivariate Gaussian distribution $\mathbf{x}_i \sim \mathcal{N}(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$, then the MLE for the parameters is given by

$$\hat{\boldsymbol{\mu}}_{mle} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i \stackrel{\text{def}}{=} \bar{\mathbf{x}}, \tag{3.1}$$

$$\hat{\boldsymbol{\Sigma}}_{mle} = \frac{1}{n} \sum_{i=1}^{n} \left(\mathbf{x}_i - \hat{\boldsymbol{\mu}}\right)\left(\mathbf{x}_i - \hat{\boldsymbol{\mu}}\right)^T = \frac{1}{n}\left(\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T\right) - \bar{\mathbf{x}}\bar{\mathbf{x}}^T. \tag{3.2}$$

That is, the MLE is just the empirical mean and empirical covariance. In the univariate case, the following familiar results are obtained:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1} x_i = \bar{x}, \tag{3.3}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i} (x_i - \bar{x})^2 = \left(\frac{1}{n}\sum_{i} x_i^2\right) - (\bar{x})^2 . \tag{3.4}$$

Proof. To prove this result, there is need for several results from matrix algebra:

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{a}}\left\{\mathbf{b}^T\mathbf{a}\right\} &= \mathbf{b}, \\
\frac{\partial}{\partial \mathbf{a}}\left\{\mathbf{a}^T\mathbf{A}\mathbf{a}\right\} &= \left(\mathbf{A} + \mathbf{A}^T\right)\mathbf{a}, \\
\frac{\partial}{\partial \mathbf{A}}\left\{\text{Trace}\,(\mathbf{BA})\right\} &= \mathbf{B}^T, \qquad (3.5)\\
\frac{\partial}{\partial \mathbf{A}}\left\{\ln\mid \mathbf{A}\mid\right\} &= \mathbf{A}^{-T} \stackrel{\text{def}}{=} \left(\mathbf{A}^{-1}\right)^T, \\
\text{Trace}\,(\mathbf{ABC}) = \text{Trace}\,(\mathbf{CAB}) &= \text{Trace}\,(\mathbf{BCA}). \qquad (3.6)
\end{aligned}$$

Equation (3.6) is called the cyclic permutation property of the trace operator [8] . Using this the scalar inner product $\mathbf{x}^T\mathbf{A}\mathbf{x}$ can be reordered as follows

$$\text{Trace}\left(\mathbf{x}^T\mathbf{A}\mathbf{x}\right) = \text{Trace}\left(\mathbf{x}\mathbf{x}^T\mathbf{A}\right) = \text{Trace}\left(\mathbf{A}\mathbf{x}\mathbf{x}^T\right). \tag{3.7}$$

Now the proof begins:
The log-likelihood is

$$\begin{aligned}
\mathcal{L}\,(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{x}) &= \ln p\,(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{x}) \\
&= \frac{n}{2}\ln\mid \boldsymbol{\Sigma}\mid - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})^T\,\boldsymbol{\Sigma}^{-1}\,(\mathbf{x}_i - \boldsymbol{\mu}) + \text{const.}
\end{aligned} \tag{3.8}$$

where $\Lambda = \boldsymbol{\Sigma}^{-1}$ is the precision matrix.

Using the substitution $\mathbf{y}_i = \mathbf{x}_i - \boldsymbol{\mu}$ and the chain rule of calculus, thus obtain

$$\frac{\partial}{\partial \boldsymbol{\mu}}(\mathbf{x}_i - \boldsymbol{\mu})^T\,\boldsymbol{\Sigma}^{-1}\,(\mathbf{x}_i - \boldsymbol{\mu}) = \frac{\partial}{\partial \mathbf{y}_i}\mathbf{y}_i^T\boldsymbol{\Sigma}^{-1}\mathbf{y}_i\frac{\partial \mathbf{y_i}}{\partial \boldsymbol{\mu}} \tag{3.9}$$

$$= -1\left(\boldsymbol{\Sigma}^{-1} + \boldsymbol{\Sigma}^{-T}\right)\mathbf{y}_i. \tag{3.10}$$

Hence

$$\frac{\partial}{\partial \boldsymbol{\mu}}\mathcal{L}\,(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = -\frac{1}{2}\sum_{i=1}^{n} -2\boldsymbol{\Sigma}^{-1}\,(\mathbf{x}_i - \boldsymbol{\mu}) = \boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu}) \tag{3.11}$$

$$= \hat{\boldsymbol{\mu}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i = \bar{\mathbf{x}}. \tag{3.12}$$

Thus the MLE estimate of the mean $\boldsymbol{\mu}$ is the empirical (sample) mean $\bar{\mathbf{x}}$.

Using the trace of a matrix [4], [9] from Equation (3.7) the log-likelihood for Λ as follows:

$$\begin{aligned}\mathcal{L}(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{x}) &= \ln p(\boldsymbol{\mu}, \boldsymbol{\Sigma} \mid \mathbf{x}) \\ &= -\frac{n}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}\sum_{i=1}^{n}(\mathbf{x}_i - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}) + \text{const} \\ &= -\frac{n}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}\sum_{i=1}^{n}\text{Trace}\left(\boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T\right) \\ &= -\frac{n}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}\text{Trace}\left(\boldsymbol{\Sigma}^{-1}\sum_{i=1}^{n}[(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T]\right) \\ &= -\frac{n}{2}\ln|\boldsymbol{\Sigma}| - \frac{1}{2}\text{Trace}\left(\mathbf{S}_\mu \boldsymbol{\Sigma}^{-1}\right),\end{aligned} \tag{3.13}$$

where

$$\mathbf{S}_\mu \stackrel{\text{def}}{=} \sum_{i=1}^{n}\left[(\mathbf{x}_i - \boldsymbol{\mu})(\mathbf{x}_i - \boldsymbol{\mu})^T\right], \tag{3.14}$$

is the scatter matrix centered on $\boldsymbol{\mu}$ [8]. Taking the derivative of Equation (3.13) w.r.t. Λ yields

$$\frac{\partial \mathcal{L}(\Lambda)}{\partial \Lambda} = \frac{n}{2}\Lambda^{-T} - \frac{1}{2}\mathbf{S}_\mu^T = 0 \tag{3.15}$$

$$\Lambda^{-T} = \Lambda^{-1} = \boldsymbol{\Sigma} = \frac{1}{n}\mathbf{S}_\mu. \tag{3.16}$$

so

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n}\left(\mathbf{x}_i - \hat{\boldsymbol{\mu}}\right)\left(\mathbf{x}_i - \hat{\boldsymbol{\mu}}\right)^T. \tag{3.17}$$

Equation (3.17) is the same as the empirical covariance matrix centered on $\boldsymbol{\mu}$. If we plug-in the MLE $\boldsymbol{\mu} = \bar{\mathbf{x}}$ (since both parameters must be simultaneously optimized), we get the standard equation for MLE of a covariance matrix. ■

3.2. The Principle of *n*-Cross Sections

Theorem 3.2. The estimates of the sufficient parameters of a multivariate Gaussian distribution can be estimated as $\widehat{\sigma} = \widehat{A}\sqrt{2\pi}$ and $\widehat{\mu} = [\widehat{\sigma}(\tau - 2\xi\eta) - 1]^{-1}$.

Proof. The multivariate Gaussian distribution with two parameters can be represented as

$$\begin{aligned} &f\,(x_1, x_2, \alpha, \beta, \rho, a, b) \\ &\quad = f\,(x_1, x_2, \mathcal{H}) \\ &\quad = A\mathrm{e}^{-\alpha(x_1-\xi a)^2-\beta(x_2-\xi b)^2-2\rho(x_1-\xi a)(x_2-\xi b)} \\ &\quad = A \exp\left[-\alpha(x_1 - \xi a)^2 - \beta(x_2 - \xi b)^2 - 2\rho(x_1 - \xi a)(x_2 - \xi b)\right] \\ &\text{where} \qquad \mathcal{H} = \{\alpha, \beta, \rho, a, b\}\,. \end{aligned} \tag{3.18}$$

The multivariate Gaussian distribution in Equation (3.18) has two variables (x_1, x_2) and six unknown parameters $\{A, \alpha, \beta, \rho, a, b\}$, which are functions of the sufficient parameters $\sigma,\ \mu$.

The unknown parameters $\mathcal{H}$ can be estimated by formulating six equations with six unknowns. This is a tedious process and instead the number of equations can be reduced by considering the proposed method by letting $x_2 = k_i x_1$, Equation (3.18) becomes a one-dimensional equation,

$$f\,(x_1, \mathcal{H}) = A \exp\left[-\alpha\,(x_1 - \xi a)^2 - \beta\,(k_i x_1 - \xi b)^2 - 2\rho\,(x_1 - \xi a)\,(k_i x_1 - \xi b)\right]. \tag{3.19}$$

Consider the argument of the exponent Λ, in Equation (3.19)

$$\Lambda = -\alpha\,(x_1 - \xi a)^2 - \beta\,(k_i x_1 - \xi b)^2 - 2\rho\,(x_1 - \xi a)\,(k_i x_1 - \xi b) \tag{3.20}$$

Simplifying Equation (3.20) yields

$$\begin{aligned} \Lambda &= -\alpha\,(x_1 - \xi a)^2 - \beta\,(k_i x_1 - \xi b)^2 - 2\rho\,(x_1 - \xi a)\,(k_i x_1 - \xi b) \\ &= \overbrace{-\alpha a^2\xi^2 + 2\alpha a\xi x_1 - \alpha x_1^2}^{-\alpha\left(x_1^2-2a\xi x_1+a^2\xi^2\right)} \overbrace{-\beta b^2\xi^2 + 2\beta bk\xi x_1 - \beta k_i^2 x_1^2}^{-\beta\left(k_i^2x_1^2-2bk_i\xi x_1+b^2\xi^2\right)} \overbrace{+2\rho b\xi x_1 - 2\rho k_i x_1^2 - 2\rho ab\xi^2 + 2\rho ak_i\xi x_1}^{-2\rho\left(k_ix_1^2-ak_i\xi x_1-b\xi x_1+ab\xi^2\right)} \\ &= x_1^2\left(-\beta k_i^2 - 2\rho k_i - \alpha\right) + 2x_1\{\alpha a\xi + \beta bk_i\xi + \rho\,(ak_i\xi + b\xi)\} \overbrace{-\xi^2\left(\alpha a^2 + \beta b^2 + 2\rho ab\right)}^{\text{Constant Term}}. \end{aligned} \tag{3.21}$$

Thus the argument of the exponent Λ, in Equation (3.19) can be expressed as

$$\begin{aligned} &\Lambda = -x_1^2\tau + 2x_1\eta - \xi^2\upsilon. \\ \text{where} \qquad &\tau = \alpha + \beta k_i^2 + 2\rho k_i,\ \eta = \alpha a\xi + \beta bk_i\xi + \rho\,(ak_i\xi + b\xi)\,, \\ \text{and} \qquad &\upsilon = \alpha a^2 + \beta b^2 + 2\rho ab. \end{aligned} \tag{3.22}$$

The first approximation of the Gaussian multivariate Equation (3.18) can now be estimated as

$$\widetilde{f}\,(x_1, \mathcal{H}) = \widetilde{A}\ \exp\left[-\tau x_1^2 + 2\xi\eta x_1\right] \qquad \text{where} \qquad \widetilde{A} = A\ \exp\left[-\xi^2\upsilon\right]. \tag{3.23}$$

Consider the argument of the exponent in Equation (3.23)

$$\phi = -\tau x_1^2 + 2\xi\eta x_1. \tag{3.24}$$

Equation (3.24) is a quadratic, thus the argument of the multivariate Gaussian has been transformed to a quadratic in x_1. By completing the square, Equation (3.24) is transformed to

$$\phi = -\tau x_1^2 + 2\xi\eta x_1 \quad = \quad -\tau\left(x_1 - \frac{\xi\eta}{\tau}\right)^2 + \frac{(\xi\eta)^2}{\tau}. \tag{3.25}$$

Letting $\zeta = \dfrac{\eta}{\tau}$ in Equation (3.25) thus

$$\phi = -\tau x_1^2 + 2\xi\eta x_1 = -\tau\,(x_1 - \xi\zeta)^2 - \xi^2\eta\zeta. \tag{3.26}$$

From the reparameterization of Equation (3.25) to Equation (3.26) the second approximation for the multivariate Gaussian Equation (3.18) is now

$$\widetilde{\widetilde{f}}\,(x_1) = \widetilde{\widetilde{A}}\,\mathrm{e}^{-\tau(x_1-\xi\zeta)^2} \qquad \text{where} \qquad \widetilde{\widetilde{A}} = \widetilde{A}\,\mathrm{e}^{-\xi^2\eta\zeta}. \tag{3.27}$$

Thus the approximations for the parameters of a two-variate Gaussian distribution with six unknowns, Equation (3.18), has been reduced to an estimation problem in one variable of three parameters. In order to estimate these parameters there is need for three equations, these equations are generated by hyperplanes that splice through the multivariate Gaussian. These hyperplanes ($\mathcal{P}_{ij}$ for $i = 0, 1, 2$ and $j = 0, 1, 2, 3, 4, 5, 6$), are separated at angles α_j and k_j. The multivariate Gaussian is then reflected on these geometric hyperplanes as one-dimensional Gaussian distributions and used to approximate the parameters of the multivariate Gaussian.

Equation (3.27) is the solution of the first order homogeneous linear differential equation with constant coefficient

$$\frac{d\widetilde{\widetilde{f}}}{dx_1} + \eta\widetilde{\widetilde{f}} = 0. \tag{3.28}$$

In order to estimate the parameters of Equation (3.27) a method of minimizing the goal function

$$\mathcal{G} = \min_{i\to n}\sum_{i=1}^{n}\left[2\tau\xi\zeta\left(\widetilde{\widetilde{f}}_i\right) - 2x_i\tau\left(\widetilde{\widetilde{f}}_i\right) - \frac{d\widetilde{\widetilde{f}}_i}{dx_i}\right]^2, \tag{3.29}$$

is applied to estimate the multivariate Gaussian on the hyperplanes. Through the method of back substitution the parameters and goal function minimization thus the estimation of the parameters for the multivariate Gaussian is done using the method of *the principle of n-cross sections*. Thus the sufficient parameters of the multivariate Gaussian distribution are estimated

1. $\widehat{\sigma} = \widehat{A}\sqrt{2\pi}$,
2. $\widehat{\mu} = \dfrac{1}{\widehat{\sigma}\,(\tau - 2\xi\eta) - 1}$.

■

4. Numerical Simulations

In this section a numerical study is done on a simulated multivariate Gaussian distribution and the parameters are estimated using the method of the principle of n-cross sections. The data generation and computations of the estimates are done using Mathematica® software.

Considering the functional form of a two-variate Gaussian distribution,

$$\begin{aligned} & f\,(x_1, x_2, \alpha, \beta, \rho, a, b) \\ & \quad = f\,(x_1, x_2, \mathcal{H}) \\ & \quad = A\mathrm{e}^{-\alpha(x_1-\xi a)^2-\beta(x_2-\xi b)^2-2\rho(x_1-\xi a)(x_2-\xi b)} \\ & \quad = A\exp\left[-\alpha(x_1-\xi a)^2-\beta(x_2-\xi b)^2-2\rho(x_1-\xi a)(x_2-\xi b)\right] \\ \text{where} \quad & \mathcal{H} = \{\alpha, \beta, \rho, a, b\}, \end{aligned} \tag{4.1}$$

is generated for $X = x_1, \ldots, x_n$, in this data set n is set to 282. Thus 282 data points are considered. Equation (4.1) has the following exact values for the parameters that are to be estimated, $A = 1.03, \xi = 0.143, \alpha = 5.203, \beta = 5.047, \rho = 3.0131, a = 0.1\, b = 0.3\, \sigma = 2.581$, and $\mu = 0.092$. Equation (4.1) with the given parameters can be represented as Figure 1.

4.1. Estimation Procedure and Results

The estimation of the multivariate Gaussian distribution is done by generating seven cross sections (hyperplanes) that intersect with the graph in Figure 1. The data of the multivariate Gaussian on the hyperplanes $\mathcal{P}_{ij}$ for $i = 0, 1, 2$ and $j = 0, 1, 2, 3, 4, 5, 6$, is generated as a two dimension data and the estimation of the parameters is done. For this numeric simulation the following hyper planes were generated:

The hyperplanes generated in Table 1 reduce the multivariate Gaussian to a one dimensional Gaussian that are used to estimate the parameters of the second approximation

$$\widetilde{\widetilde{f}}\,(x_1) = \widetilde{\widetilde{A}}\,\mathrm{e}^{-\tau(x_1-\xi\zeta)^2} \qquad \text{where} \qquad \widetilde{\widetilde{A}} = \widetilde{A}\,\mathrm{e}^{-\xi^2\eta\zeta}. \tag{4.2}$$

The unknown parameters $\widetilde{\widetilde{A}}$, τ and ζ are approximated by minimizing the goal function

$$\mathcal{G} = \min_{i\to n}\sum_{i=1}^{n}\left[2\tau\xi\zeta\left(\widetilde{\widetilde{f}}_i\right) - 2x_i\tau\left(\widetilde{\widetilde{f}}_i\right) - \frac{d\widetilde{\widetilde{f}}_i}{dx_i}\right]^2, \tag{4.3}$$

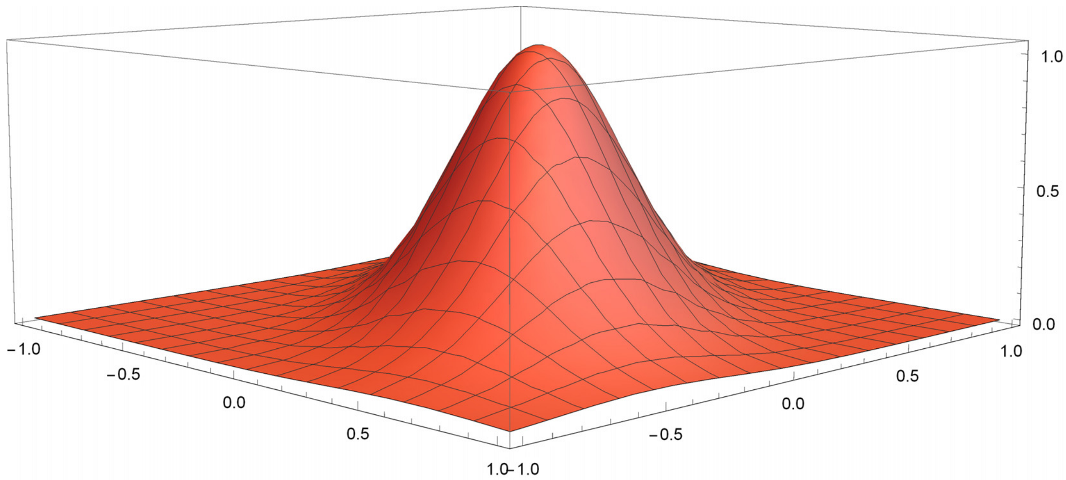

Figure 1: The Multivariate Gaussian Density Function for exact values $A = 1.03$, $\xi = 0.143$, $\alpha = 5.203$, $\beta = 5.047$, $\rho = 3.0131$, $a = 0.1$ and $b = 0.3$

Table 1: Hyperplanes for the Numeric Simulation

Hyperplane $\mathcal{P}_{ij}$	k_j	Angle α_j
$\mathcal{P}_{00}$	$k_0 = 0$	$\alpha_0 = 0$
$\mathcal{P}_{11}$	$k_1 = \sqrt{3}$	$\alpha_1 = \frac{\pi}{3}$
$\mathcal{P}_{22}$	$k_2 = -\sqrt{3}$	$\alpha_2 = -\frac{\pi}{3}$
$\mathcal{P}_{13}$	$k_3 = 1$	$\alpha_3 = \frac{\pi}{4}$
$\mathcal{P}_{24}$	$k_4 = -1$	$\alpha_4 = -\frac{\pi}{4}$
$\mathcal{P}_{15}$	$k_5 = \frac{\sqrt{2}}{2}$	$\alpha_5 = 0.6154797$
$\mathcal{P}_{26}$	$k_6 = -\frac{\sqrt{2}}{2}$	$\alpha_6 = -0.6154797$

The following results are obtained for the estimation of the second approximation using the generated hyperplanes $(\mathcal{P}_{ij})$ assuming $\xi \lll$ thus $\xi = 1 \times 10^{-1}$. Care should be taken when approximating the values of the parameter functions from the system of equations as the absolute values are considered.

Since

$$\tau = \alpha + \beta k_j^2 + 2\rho, \tag{4.4}$$

$$\eta = \alpha a\xi + \beta b k_j \xi + \rho\left(a k_j \xi + b\xi\right), \tag{4.5}$$

$$\upsilon = \alpha a^2 + \beta b^2 + 2\rho ab. \tag{4.6}$$

From Table 2 a system of equations for estimating the unknown parameters $\widetilde{\widetilde{A}}$, τ, ζ and η

Table 2: Minimization of the Goal Function for the Second Approximation using the Hyperplanes $\mathcal{P}_{ij}$

Hyperplane	$\widetilde{\widetilde{A}}$	τ	ζ	η
$\mathcal{P}_{00}$	1.022595555	5.202012751	0.037119404	0.193095613
$\mathcal{P}_{11}$	1.018863171	30.75061713	0.019198013	0.59035074
$\mathcal{P}_{22}$	1.02193435	9.903096606	$\|-0.026856134\|$	$\|-0.265958891\|$
$\mathcal{P}_{13}$	1.024775658	16.26751179	0.0264374	0.430070721
$\mathcal{P}_{24}$	1.016342564	4.222994171	$\|-0.015267556\|$	$\|-0.064474802\|$
$\mathcal{P}_{15}$	1.025458907	11.98295359	0.030277166	0.362809873
$\mathcal{P}_{26}$	1.015285857	3.464224689	0.003783799	0.013107929

can be generated, Equations (4.4, 4.5, and 4.6).

Table 3 represents the estimated unknown parameters in the functional form of the multivariate Gaussian distribution obtained using the principle of n-cross sections. Using appropriate algebraic formulations the required sufficient parameters σ and μ are then computed as

$$\widehat{\sigma} = \widehat{A}\sqrt{2\pi}, \qquad \text{and} \tag{4.7}$$

$$\widehat{\mu} = \frac{1}{\widehat{\sigma}\,(\tau - 2\xi\eta) - 1}. \tag{4.8}$$

Table 3: Comparison of the Exact solutions and Parameter Estimation using the principle of n-cross sections

	Values			Percentage Errors	
Parameter	Exact	PnCS	MLE	PnCS	MLE
A	1.030	1.023	0.632	0.680	38.641
ξ	0.143	0.100	0.180	30.070	25.874
α	5.203	5.426	5.720	4.286	9.937
β	5.047	5.355	5.551	6.103	9.986
ρ	3.013	2.744	3.500	8.928	16.163
a	0.100	0.151	0.028	51.000	72.000
b	0.300	0.437	0.253	45.667	15.667
σ	3.000	2.582	1.175	13.933	60.833
μ	0.100	0.092	0.035	8.000	65.000

Figure 2 represents the approximated multivariate Gaussian distribution of Equation (4.1) represented by Figure 1.

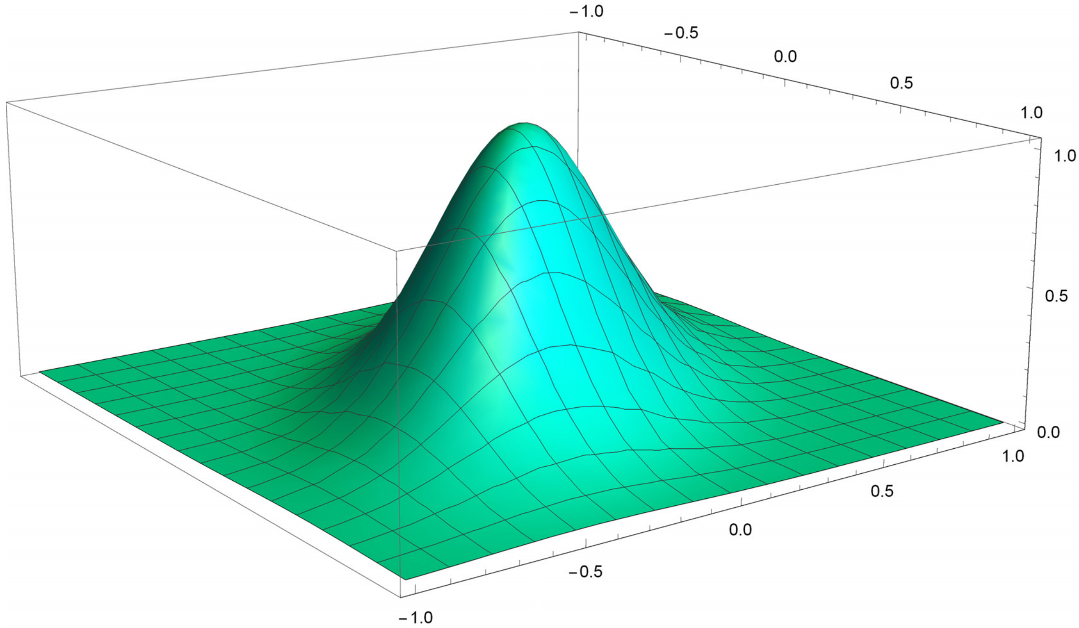

Figure 2: The Multivariate Gaussian Density Function for Approximated values $A = 1.023$, $\xi = 0.100$, $\alpha = 5.4256$, $\beta = 5.3548$, $\rho = 2.74356$, $a = 0.151$ and $b = 0.437$

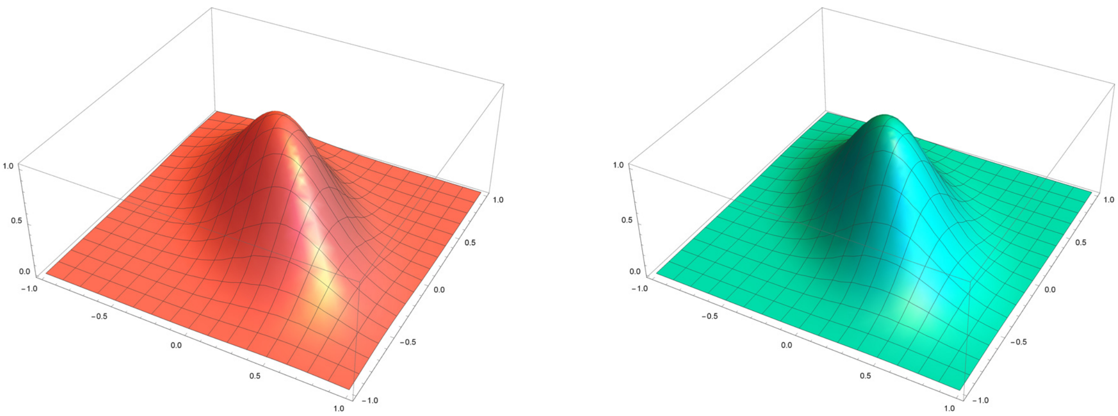

Figure 3: The Multivariate Gaussian Density Function for Exact Values Compared with the Approximated Values

5. Conclusion

The approximation of the multivariate Gaussian distribution is done using the proposed approach, principle n-cross sections (PnCS). The developed method estimates the sufficient parameters of the multivariate Gaussian distribution very close to the exact values Table 3. It is also observed that the PnCS method performs better than the maximum likelihood estimator (MLE) method, as seen from accumulated error margin of 18.74% and 34.90% respectively.

In Figure 3, a comparison of the exact and approximated multivariate Gaussian distributions is done. Hence the PnCS could be a useful estimation approach than the MLE.

References

[1] Q.P. Li. *Speaker Authentication*. Springer, 2012.

[2] A.C. Rencher. *Methods of Multivariate Analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons, INC. Publication, second edition, 2002.

[3] T.D. Wickens. *Elementary Signal Detection Theory*. Oxford UniversitOxford , Inc., 2002.

[4] N.H. Timm. *Applied Multivariate Analysis*. Springer Texts in Statistics. Springer-Verlag New York, Inc., 2002.

[5] R.D. Fricker, Jr. *Statistical Methods in Counterterrorism Game Theory*. Springer, 2006.

[6] C. Chatfield and A.J. Collins. *Introduction To Multivariate Analysis*. Springer-Science+Bussiness Media, B.V., 1980.

[7] D.W. Scott and S.R. Sain. Multi-dimensional density estimation. *Preprint submitted to Elsevier Science*, 2004.

[8] K.P. Murphy. *Machine Learning: A probabilistic Perspective*. The MIT Press, 2012.

[9] A.R. Gallant. *Nonlinear Statistical Models*, volume Wiely Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc., 1987.

[10] S.R. Searle. *Linear Models*. John Wiley & Sons, Inc., 1971.

[11] S. Kotz, N. Balakrishnan, and N.L. Johnson. *Continuous Multivariate Distributions*. John Wiley & Sons, Inc., 2000.

[12] B. Fichet, D. Poccolo, R. Verde, and M. Vinci. *Classification and Multivariate Analysis for Complex Data Structures*. Springer-Verlag Berlin Heidelberg, 2011.

[13] S. Kotz, T.J. Kozubowski, and K. Podgorski. *The Laplace Distribution and Generalizations*. Springer Science+Bussiness Media, LLC, 2001.