

A note on Gaussian distributions in \mathbb{R}^n

B G MANJUNATH and K R PARTHASARATHY

Indian Statistical Institute, Delhi Centre, 7, S. J. S. Sansanwal Marg,
New Delhi 110 016, India
E-mail: bgmanjunath@gmail.com; krp@isid.ac.in

MS received 11 August 2011; revised 19 October 2011

Abstract. Given any finite set \mathcal{F} of $(n - 1)$ -dimensional subspaces of \mathbb{R}^n we give examples of nonGaussian probability measures in \mathbb{R}^n whose marginal distribution in each subspace from \mathcal{F} is Gaussian. However, if \mathcal{F} is an infinite family of such $(n - 1)$ -dimensional subspaces then such a nonGaussian probability measure in \mathbb{R}^n does not exist.

Keywords. Gaussian distribution; characteristic function; homogeneous polynomial; linear functionals; nonunimodality; Hermite polynomial.

1. Introduction

Starting with the simple example of E Nelson as cited by Feller in [1] we have from [2–4], etc., as well as §10 of [5], several examples of bivariate and multivariate nonGaussian distributions under which many linear functionals can have a Gaussian distribution on the real line. These results suggest the possibility of characterizing a Gaussian distribution in \mathbb{R}^n through properties of classes of linear functionals. Motivated by Nelson's example in [1] and the bivariate construction in [2] we introduce a perturbation of the standard Gaussian density function in \mathbb{R}^n exhibiting the following interesting features:

- (1) Given any finite set $\{S_j, 1 \leq j \leq N\}$ of $(n - 1)$ -dimensional subspaces it has a marginal density function which is the standard Gaussian in each S_j , $j \in \{1, 2, \dots, N\}$;
- (2) There can exist linear functionals whose distributions may have nonunimodal density functions;
- (3) For certain choices of subspaces the nonGaussian perturbation can be so chosen that any real symmetric measurable function of all the n co-ordinates has its distribution preserved. In particular, the sum of squares of all the co-ordinates can have the χ^2 distribution with n degrees of freedom.

We also demonstrate the following characterization of the multivariate Gaussian distribution. Suppose $\{S_j, j = 1, 2, \dots\}$ is a countably infinite set of $(n - 1)$ -dimensional subspaces of \mathbb{R}^n and μ is a probability measure in \mathbb{R}^n such that the projection of μ in each subspace S_j is Gaussian. Then μ itself is Gaussian. This is a generalization of the characterization in [2] and a more precise version of the result in [4].

Our proofs follow the steps in [2] and we use some additional geometric and topological arguments of a very elementary kind.

2. A perturbation of the Gaussian characteristic function

We begin by examining a small perturbation of the characteristic function of the n -variate standard Gaussian distribution with mean vector $\mathbf{0}$ and covariance matrix \mathbf{I} as follows. Choose and fix any homogeneous polynomial \mathcal{P} of even degree $2k$ in n real variables t_1, t_2, \dots, t_n and define

$$\Phi(\mathbf{t}; \varepsilon, \sigma, \mathcal{P}) = e^{-\frac{1}{2}|\mathbf{t}|^2} + \varepsilon e^{-\frac{1}{2}\sigma^2|\mathbf{t}|^2} \mathcal{P}(\mathbf{t}), \quad \mathbf{t} \in \mathbb{R}^n \quad (2.1)$$

where $\mathbf{t} = (t_1, \dots, t_n)^T$, ε is a real parameter and σ is a parameter satisfying $0 < \sigma < 1$. Here

$$|\mathbf{t}|^2 = (t_1^2 + \dots + t_n^2).$$

Clearly, $\Phi(\cdot; \varepsilon, \sigma, \mathcal{P})$ is a real analytic function on \mathbb{R}^n satisfying

$$\begin{aligned} \Phi(\mathbf{0}; \varepsilon, \sigma, \mathcal{P}) &= 1, \\ \Phi(-\mathbf{t}; \varepsilon, \sigma, \mathcal{P}) &= \Phi(\mathbf{t}; \varepsilon, \sigma, \mathcal{P}). \end{aligned} \quad (2.2)$$

Let

$$Z_{\mathcal{P}} = \{\mathbf{t} | \mathcal{P}(\mathbf{t}) = 0, \quad \mathbf{t} \in \mathbb{R}^n\} \quad (2.3)$$

be the set of zeros of \mathcal{P} in \mathbb{R}^n .

Since we are interested in the inverse Fourier transform of Φ we introduce the renormalized polynomial \mathfrak{P} in the form of a formal definition.

DEFINITION 2.1

Let

$$\mathfrak{N}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

and let $H_m(x)$ be the m -th Hermite polynomial defined by

$$\frac{d^m}{dx^m} \mathfrak{N}(x) = (-1)^m H_m(x) \mathfrak{N}(x), \quad m = 0, 1, 2, \dots$$

(as in [1]). For any real polynomial \mathcal{P} in n real variables given by

$$\mathcal{P}(t_1, t_2, \dots, t_n) = \sum_{\mathbf{m}} a_{m_1, m_2, \dots, m_n} t_1^{m_1} t_2^{m_2} \dots t_n^{m_n}$$

its renormalized version \mathcal{P} is defined by

$$\mathcal{P} : (x_1, \dots, x_n) = \sum_{\mathbf{m}} a_{m_1, m_2, \dots, m_n} H_{m_1}(x_1) H_{m_2}(x_2) \dots H_{m_n}(x_n).$$

Note that for a homogeneous polynomial, its renormalized version need not be homogeneous.

Since the function Φ in (2.1) is in $\mathbb{L}_1(\mathbb{R}^n)$ its inverse Fourier transform f is defined by

$$\begin{aligned} f(\mathbf{x}; \varepsilon, \sigma, \mathcal{P}) &= \frac{1}{(2\pi)^n} \int e^{-it^T \mathbf{x}} \Phi(\mathbf{t}; \varepsilon, \sigma, \mathcal{P}) dt_1 dt_2 \dots dt_n \\ &= \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}|\mathbf{x}|^2} \\ &\quad + \varepsilon \frac{1}{(2\pi)^n} \int e^{-it^T \mathbf{x}} e^{-\frac{1}{2}\sigma^2|\mathbf{t}|^2} \mathcal{P}(\mathbf{t}) dt_1 \dots dt_n. \end{aligned} \quad (2.4)$$

First, we note that

$$\frac{1}{(2\pi)^n} \int e^{-it^T \mathbf{x}} e^{-\frac{1}{2}\sigma^2|\mathbf{t}|^2} dt_1 dt_2 \dots dt_n = \frac{1}{\sigma^n} \prod_{j=1}^n \mathfrak{N}\left(\frac{x_j}{\sigma}\right).$$

Repeated differentiation with respect to x_1, x_2, \dots, x_n shows that for the homogeneous polynomial \mathcal{P} of degree $2k$ we have

$$\begin{aligned} &\frac{1}{(2\pi)^n} \int e^{-it^T \mathbf{x}} e^{-\frac{1}{2}\sigma^2|\mathbf{t}|^2} \mathcal{P}(\mathbf{t}) dt_1 dt_2 \dots dt_n \\ &= \frac{1}{\sigma^n} \mathcal{P}\left(i \frac{\partial}{\partial x_1}, \dots, i \frac{\partial}{\partial x_n}\right) \left\{ \prod_{j=1}^n \mathfrak{N}\left(\frac{x_j}{\sigma}\right) \right\} \\ &= \frac{(-1)^k}{\sigma^{n+2k}} : \mathcal{P} : \left(\frac{x_1}{\sigma}, \dots, \frac{x_n}{\sigma}\right) \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2\sigma^2}|\mathbf{x}|^2}. \end{aligned}$$

Thus the inverse Fourier transform (2.4) assumes the form

$$\begin{aligned} f(\mathbf{x}; \varepsilon, \sigma, \mathcal{P}) &= \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}|\mathbf{x}|^2} \left\{ 1 + \frac{(-1)^k \varepsilon}{\sigma^{n+2k}} : \mathcal{P} : \left(\frac{x_1}{\sigma}, \dots, \frac{x_n}{\sigma}\right) e^{-\frac{1}{2\sigma^2}|\mathbf{x}|^2(1-\sigma^2)} \right\}. \end{aligned} \quad (2.5)$$

Since, by assumption, $1 - \sigma^2 > 0$ the positive constant $K(\sigma, \mathcal{P})$ defined by

$$K(\sigma, \mathcal{P}) = \sup_{\mathbf{x} \in \mathbb{R}^n} \frac{|\mathcal{P} : (x_1, \dots, x_n)|}{\sigma^{n+2k}} e^{-\frac{1}{2}|\mathbf{x}|^2(1-\sigma^2)} \quad (2.6)$$

is finite and for all $\mathbf{x} \in \mathbb{R}^n$,

$$f(\mathbf{x}; \varepsilon, \sigma, \mathcal{P}) \geq 0, \quad \text{if } |\varepsilon| \leq K^{-1}(\sigma, \mathcal{P}).$$

We observe that $\Phi(\cdot; \varepsilon, \sigma, \mathcal{P})$ is a real characteristic function of the probability density function $f(\cdot; \varepsilon, \sigma, \mathcal{P})$ defined by (2.5) for any $\varepsilon \in [-K^{-1}(\sigma, \mathcal{P}), K^{-1}(\sigma, \mathcal{P})]$. Here we have made use of property (2.2). Thus we can summarize the discussion above as a theorem.

Theorem 2.2. Let $0 < \sigma < 1$, \mathcal{P} be a real homogeneous polynomial in n variables of even degree $2k$, $K(\sigma, \mathcal{P})$ the positive constant defined by (2.6) and $\varepsilon \in [-K^{-1}(\sigma, \mathcal{P}), K^{-1}(\sigma, \mathcal{P})]$. Then the function $\Phi(\cdot; \varepsilon, \sigma, \mathcal{P})$ defined by (2.1) is the characteristic function of a probability density function $f(\cdot; \varepsilon, \sigma, \mathcal{P})$ defined by (2.5). Under this density function $f(\cdot; \varepsilon, \sigma, \mathcal{P})$ the linear functional $\mathbf{x} \mapsto \mathbf{a}^T \mathbf{x}$ with $|\mathbf{a}| = 1$ has characteristic function $\varphi_{\mathbf{a}}$ and probability density function $g_{\mathbf{a}}$ on the real line given respectively by

$$\varphi_{\mathbf{a}}(t) = e^{-\frac{1}{2}t^2} + \varepsilon \mathcal{P}(\mathbf{a}) e^{-\frac{1}{2}\sigma^2 t^2} t^{2k}, \quad t \in \mathbb{R} \quad (2.7)$$

$$g_{\mathbf{a}}(x) = \frac{1}{\sqrt{2\pi}} \left\{ e^{-\frac{1}{2}x^2} + \frac{(-1)^k \varepsilon \mathcal{P}(\mathbf{a})}{\sigma^{2k+1}} H_{2k} \left(\frac{x}{\sigma} \right) e^{-\frac{1}{2\sigma^2}x^2} \right\}. \quad (2.8)$$

In particular, for any $\mathbf{a} \in Z_{\mathcal{P}}$, the linear functional $\mathbf{a}^T \mathbf{x}$ has the normal distribution with mean 0 and variance $|\mathbf{a}|^2$ but $f(\cdot; \varepsilon, \sigma, \mathcal{P})$ is a nonGaussian density function for any $\varepsilon \in [-K^{-1}(\sigma, \mathcal{P}), K^{-1}(\sigma, \mathcal{P})] \setminus \{0\}$.

Proof. The first part is immediate from the discussion preceding the statement of the theorem. To prove the second part, we note that the characteristic function $\varphi_{\mathbf{a}}(t)$ of the linear functional $\mathbf{a}^T \mathbf{x}$ under the density function $f(\cdot; \varepsilon, \sigma, \mathcal{P})$ is $\Phi(t\mathbf{a}; \varepsilon, \sigma, \mathcal{P})$ and (2.7) follows from (2.1) and the homogeneity of \mathcal{P} . Now (2.8) follows from Fourier inversion of (2.7). If $0 \neq \mathbf{a} \in Z_{\mathcal{P}}$, then $0 = \mathcal{P}(\mathbf{a}) = \mathcal{P}\left(\frac{\mathbf{a}}{|\mathbf{a}|}\right)$ and therefore

$$\varphi_{\frac{\mathbf{a}}{|\mathbf{a}|}}(t) = e^{-\frac{1}{2}t^2}.$$

Hence $\mathbf{a}^T \mathbf{x}$ is normally distributed with mean 0 and variance $|\mathbf{a}|^2$. □

COROLLARY 2.3

Let $\{S_j, 1 \leq j \leq N\}$ be any finite set of $(n-1)$ -dimensional subspaces of \mathbb{R}^n . Then there exists a nonGaussian analytic probability density function whose projection on S_j is Gaussian for each $j \in \{1, 2, \dots, N\}$.

Proof. By adding one more $(n-1)$ -dimensional subspace to the collection $\{S_j, 1 \leq j \leq N\}$, if necessary, we may assume without loss of generality that N is even. Choose a unit vector $\mathbf{a}^{(j)} \in S_j^\perp$ for each j and define the homogeneous real polynomial \mathcal{P} of degree N by

$$\mathcal{P}(\mathbf{t}) = \prod_{j=1}^N \mathbf{a}^{(j)T} \mathbf{t}, \quad \mathbf{t} \in \mathbb{R}^n.$$

Clearly,

$$\mathcal{P}(\mathbf{t}) = 0, \quad \text{if } \mathbf{t} \in \bigcup_{j=1}^N S_j.$$

In other words,

$$\bigcup_{j=1}^N S_j \subset Z_{\mathcal{P}}.$$

If we choose μ to be the probability measure with the density function $f(\cdot; \varepsilon, \sigma, \mathcal{P})$, $0 \neq \varepsilon$ in $[-K^{-1}(\sigma, \mathcal{P}), K^{-1}(\sigma, \mathcal{P})]$ in Theorem 2.2 it follows immediately from the last part of the theorem that every linear functional of the form $\mathbf{b}^T \mathbf{x}$ has a normal distribution with mean 0 and variance $|\mathbf{b}|^2$ whenever $\mathbf{b} \in Z_{\mathcal{P}}$. This completes the proof. \square

Remark 2.4. In the context of understanding the modes of the density function $g_{\mathbf{a}}(x)$ given by (2.8), it is of interest to note that

$$\begin{aligned} & \{x | x \neq 0, g'_{\mathbf{a}}(x) = 0\} \\ &= \left\{ x \mid x \neq 0, e^{\frac{x^2}{2}(\frac{1}{\sigma^2}-1)} + \frac{(-1)^k \varepsilon \mathcal{P}(\mathbf{a})}{\sigma^{2k+2}} \frac{H_{2k+1}(\frac{x}{\sigma})}{x} = 0 \right\}. \end{aligned}$$

Indeed, this is obtained by straightforward differentiation and using the recurrence relation $H_{2k+1}(x) = x H_{2k}(x) - H'_{2k}(x)$.

Example 2.5. Let n be even and

$$\begin{aligned} \mathcal{P}(t_1, t_2, \dots, t_n) &= t_1 t_2 \dots t_n \prod_{i>j} (t_i^2 - t_j^2) \\ &= t_1 t_2 \dots t_n \begin{vmatrix} 1 & 1 & \dots & 1 \\ t_1^2 & t_2^2 & \dots & t_n^2 \\ t_1^4 & t_2^4 & \dots & t_n^4 \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ t_1^{2(n-1)} & t_2^{2(n-1)} & \dots & t_n^{2(n-1)} \end{vmatrix}. \end{aligned} \quad (2.9)$$

Then \mathcal{P} is a polynomial of even degree n^2 , which is antisymmetric in the variables t_1, t_2, \dots, t_n . The renormalized version $:\mathcal{P}:$ of \mathcal{P} is given by

$$:\mathcal{P}:(x_1, x_2, \dots, x_n) = \begin{vmatrix} H_1(x_1) & H_1(x_2) & \dots & H_1(x_n) \\ H_3(x_1) & H_3(x_2) & \dots & H_3(x_n) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ H_{2n+1}(x_1) & H_{2n+1}(x_2) & \dots & H_{2n+1}(x_n) \end{vmatrix}. \quad (2.10)$$

In particular, $:\mathcal{P}:$ is antisymmetric in the variables x_1, x_2, \dots, x_n . Fixing $0 < \sigma < 1$ we get for each $\varepsilon \in [-K^{-1}(\sigma, \mathcal{P}), K^{-1}(\sigma, \mathcal{P})]$, with $K(\sigma, \mathcal{P})$ being determined

by (2.6), (2.10) and $k = \frac{1}{2}n^2$, the probability density function $f(\cdot; \varepsilon, \sigma, \mathcal{P})$ is given by

$$f(\mathbf{x}; \varepsilon, \sigma, \mathcal{P}) = \frac{1}{(\sqrt{2\pi})^n} \left\{ e^{-\frac{1}{2}|\mathbf{x}|^2} + \frac{\varepsilon}{\sigma^{n(n+1)}} \begin{vmatrix} H_1\left(\frac{x_1}{\sigma}\right) & H_1\left(\frac{x_2}{\sigma}\right) & \dots & H_1\left(\frac{x_n}{\sigma}\right) \\ H_3\left(\frac{x_1}{\sigma}\right) & H_3\left(\frac{x_2}{\sigma}\right) & \dots & H_3\left(\frac{x_n}{\sigma}\right) \\ \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \dots & \cdot \\ H_{2n+1}\left(\frac{x_1}{\sigma}\right) & H_{2n+1}\left(\frac{x_2}{\sigma}\right) & \dots & H_{2n+1}\left(\frac{x_n}{\sigma}\right) \end{vmatrix} e^{-\frac{1}{2}\sigma^2|\mathbf{x}|^2} \right\}. \quad (2.11)$$

By Theorem 2.2 and its corollary we conclude that the projection of this density function on the $(n-1)$ -dimensional hyperplanes $\{\mathbf{x}|x_j = 0\}$, $1 \leq j \leq n$; $\{\mathbf{x}|x_i - x_j = 0\}$, $1 \leq i \leq j \leq n$ and $\{\mathbf{x}|x_i + x_j = 0\}$, $1 \leq i \leq j \leq n$ are all $(n-1)$ -dimensional Gaussian densities.

If $g(x_1, x_2, \dots, x_n)$ is any bounded continuous function which is symmetric in the variables x_1, x_2, \dots, x_n then the function $g : \mathcal{P}$ is an antisymmetric function in \mathbb{R}^n and therefore

$$\int_{\mathbb{R}^n} (g : \mathcal{P})(x_1, x_2, \dots, x_n) e^{-\frac{1}{2\sigma^2}|\mathbf{x}|^2} dx_1 dx_2 \dots dx_n = 0.$$

Thus

$$\begin{aligned} & \int_{\mathbb{R}^n} g(x_1, \dots, x_n) f(\mathbf{x}; \varepsilon, \sigma, \mathcal{P}) dx_1 dx_2 \dots dx_n \\ &= \int_{\mathbb{R}^n} g(x_1, x_2, \dots, x_n) \frac{1}{(\sqrt{2\pi})^n} e^{-\frac{1}{2}|\mathbf{x}|^2} dx_1 dx_2 \dots dx_n. \end{aligned}$$

In other words, for $0 \neq \varepsilon \in [-K^{-1}(\sigma, \mathcal{P}), K^{-1}(\sigma, \mathcal{P})]$, any symmetric measurable function g on \mathbb{R}^n has the property that its distribution under the nonGaussian density function $f(\mathbf{x}; \varepsilon, \sigma, \mathcal{P})$ in (2.11) is the same as its distribution under the standard Gaussian density function with mean $\mathbf{0}$ and covariance matrix \mathbf{I} .

Example 2.6. We now specialize Example 2.5 to the case $n = 2$, $\sigma = 2^{-1/2}$ when

$$\begin{aligned} \mathcal{P}(t_1, t_2) &= t_1 t_2 (t_1^2 - t_2^2), \\ : \mathcal{P} : (x_1, x_2) &= H_3(x_1)H_1(x_2) - H_1(x_1)H_3(x_2) \\ &= x_1^3 x_2 - x_2^3 x_1. \end{aligned}$$

A simple computation shows that

$$\begin{aligned} K(\sigma, \mathcal{P}) &= 8 \sup |x_1^3 x_2 - x_2^3 x_1| e^{-\frac{1}{4}(x_1^2 + x_2^2)} \\ &= 128e^{-2}. \end{aligned}$$

This supremum is easily evaluated by switching over to the polar co-ordinates $x_1 = r \cos \theta$, $x_2 = r \sin \theta$. Then

$$f(\mathbf{x}; \varepsilon, \sigma, \mathcal{P}) = \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)} \{1 + 32\varepsilon(x_1^3 x_2 - x_2^3 x_1) e^{-\frac{1}{2}(x_1^2 + x_2^2)}\} \quad (2.12)$$

which is a probability density function whenever

$$|\varepsilon| \leq \frac{e^2}{128}.$$

At $\varepsilon = 0$, it is the standard normal density function with mean $\mathbf{0}$ and covariance matrix \mathbf{I} . We write $\eta = 32\varepsilon$ and express the density function (2.12) as

$$f_\eta(x_1, x_2) = \frac{1}{2\pi} e^{-\frac{1}{2}(x_1^2 + x_2^2)} \{1 + \eta(x_1^3 x_2 - x_2^3 x_1) e^{-\frac{1}{2}(x_1^2 + x_2^2)}\} \quad (2.13)$$

where (see figure 1.)

$$|\eta| \leq \frac{e^2}{4}.$$

When $\mathbf{a} = (\sin \theta, \cos \theta)$ the density function g_θ of the linear functional $\mathbf{x} \mapsto x_1 \sin \theta + x_2 \cos \theta$, under f_η is given by the formula (2.8) of Theorem 2.2 as

$$g_\theta(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \left\{ 1 - \frac{\sqrt{2} \eta \sin(4\theta)}{32} (4x^4 - 12x^2 + 3) e^{-\frac{1}{2}x^2} \right\}. \quad (2.14)$$

Thus

$$g'_\theta(x) = \frac{-x}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \left\{ e^{\frac{1}{2}x^2} - \frac{\sqrt{2} \eta \sin(4\theta)}{16} (4x^4 - 20x^2 + 15) e^{-\frac{1}{2}x^2} \right\}.$$

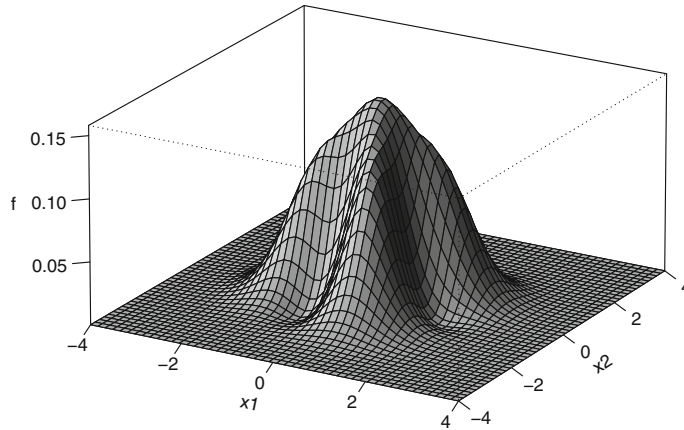


Figure 1. Bivariate density $f_\eta(x_1, x_2)$ at $\eta = e^2/4$.

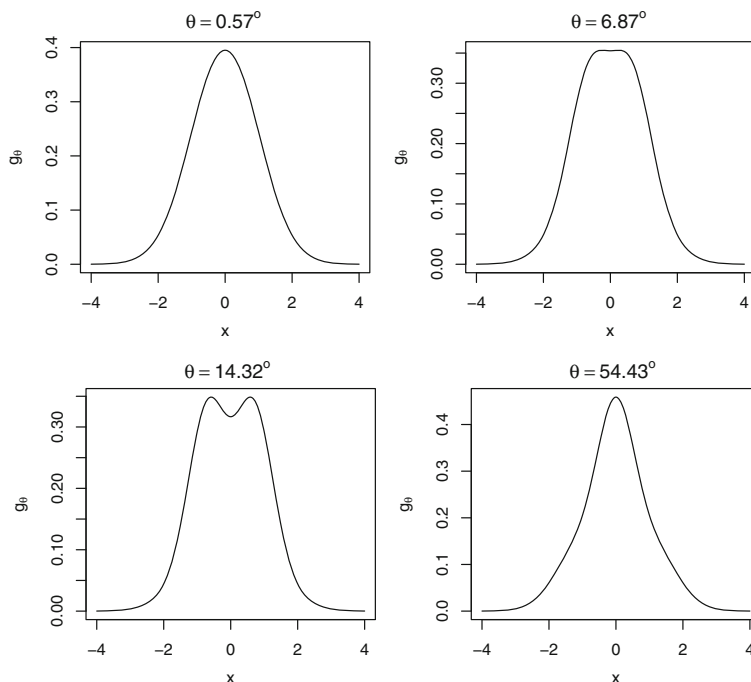


Figure 2. Nonunimodality of g_θ .

It is not difficult to find values of η in the range $(0, \frac{1}{4}e^2]$ and angle θ for which

$$\left\{ x \mid e^{\frac{1}{2}x^2} - \frac{\sqrt{2} \eta \sin(4\theta)}{16} (4x^4 - 20x^2 + 15) = 0, \quad x \neq 0 \right\} \neq \emptyset. \quad (2.15)$$

This reveals the possibility of nonunimodality of the density of some linear functionals under the joint density f_η . For an illustration, see figure 2.

3. A characterization of Gaussian distributions in \mathbb{R}^n

In the context of the Corollary to Theorem 2.2 we have the following characterization of a Gaussian distribution in \mathbb{R}^n when the number N of $(n-1)$ -dimensional subspaces in the corollary is countably infinite.

Theorem 3.1. *Let $\{S_j, j = 1, 2, \dots\}$ be a countably infinite set of $(n-1)$ -dimensional subspaces of \mathbb{R}^n and let μ be a probability measure in \mathbb{R}^n whose projection on S_j is Gaussian for each $j = 1, 2, \dots$. Then μ is Gaussian.*

Proof. The fact that the projection of μ on the two distinct $(n-1)$ -dimensional subspaces S_1 and S_2 are Gaussian implies that the multivariate Laplace transform $\hat{\mu}$ of μ given by

$$\hat{\mu}(z_1, \dots, z_n) = \int \exp(z_1 x_1 + \dots + z_n x_n) \mu(dx_1 dx_2 \dots dx_n) \quad (3.1)$$

is well-defined for $\mathbf{z} \in \mathbb{C}^n$ and analytic in each of the complex variables z_j , $j = 1, \dots, n$. Let \mathbf{m} and Σ be respectively the mean vector and covariance matrix of the \mathbb{R}^n valued random variable \mathbf{x} with distribution μ .

Choose and fix a unit vector $\mathbf{a}^{(j)} \in S_j^\perp$ for each $j = 1, 2, \dots$. Suppose

$$\begin{aligned}\mathbf{a}^{(j)T} &= (a_{j1}, \dots, a_{jn}), \quad j = 1, 2, \dots, \\ \alpha_j &= \max_{1 \leq r \leq n} |a_{jr}|.\end{aligned}$$

Since

$$\sum_{r=1}^n a_{jr}^2 = 1, \quad \forall j$$

it follows that $\alpha_j \geq n^{-1/2}$, $\forall j$. There exists an r_0 such that $a_{jr_0} = \alpha_j$ for infinitely many values of j . Restricting ourselves to this infinite set of j 's and assuming $r_0 = 1$ without loss of generality we may as well assume that

$$\begin{aligned}\mathbf{a}^{(j)} &= (a_{j1}, \dots, a_{jn})^T, \\ |a_{j1}| &= \max_{1 \leq r \leq n} |a_{jr}|, \quad \forall j = 1, 2, \dots, \\ |a_{j1}| &\geq n^{-1/2}, \quad \forall j.\end{aligned}$$

Now consider the $(n-1)$ -dimensional vector $\mathbf{b}^{(j)}$ defined by

$$\mathbf{b}^{(j)T} = \left(\frac{a_{j2}}{a_{j1}}, \frac{a_{j3}}{a_{j1}}, \dots, \frac{a_{jn}}{a_{j1}} \right), \quad j = 1, 2, \dots$$

where

$$\left| \frac{a_{jr}}{a_{j1}} \right| \leq 1, \quad \forall r = 2, 3, \dots, n.$$

Thus all the vectors $\mathbf{b}^{(j)}$ are distinct and they constitute a bounded countable set in $\mathbb{R}^{(n-1)}$. Define the set

$$\mathbb{D} = \bigcap_{j < i} \left\{ \mathbf{s} \mid \mathbf{s} \in \mathbb{R}^{(n-1)}, (\mathbf{b}^{(j)} - \mathbf{b}^{(i)})^T \mathbf{s} \neq 0 \right\}.$$

Being a countable intersection of dense open sets it follows from the Baire category theorem that \mathbb{D} is dense in $\mathbb{R}^{(n-1)}$. Let now

$$\mathbf{s} = (s_2, s_3, \dots, s_n)^T \in \mathbb{R}^{(n-1)}$$

be any fixed point in \mathbb{D} . Define

$$s_{j1} = -\mathbf{b}^{(j)T} \mathbf{s}, \quad j = 1, 2, \dots$$

By the definition of \mathbb{D} , $\{s_{j1}, j = 1, 2, \dots\}$ is a bounded and countably infinite set of distinct points on the real line. Furthermore,

$$a_{j1}s_{j1} + a_{j2}s_2 + \dots + a_{jn}s_n = 0, \quad \forall j.$$

In other words, $(s_{j1}, s_2, \dots, s_n)^T \in S_j$ for each j . By hypothesis the linear functional $s_{j1}x_1 + s_2x_2 + \dots + s_nx_n$ has a normal distribution with mean $s_{j1}m_1 + s_2m_2 + \dots + s_nm_n$ and variance $(s_{j1}, s_2, \dots, s_n)\Sigma(s_{j1}, s_2, \dots, s_n)^T$. Defining

$$\psi(z_1, \dots, z_n) = \exp\left(\mathbf{m}^T \mathbf{z} + \frac{1}{2} \mathbf{z}^T \Sigma \mathbf{z}\right), \quad \mathbf{z} \in \mathbb{C}^n$$

we conclude that the Laplace transform $\hat{\mu}$ defined by (3.1) and the function ψ satisfy the relation

$$\hat{\mu}(s_{j1}, s_2, \dots, s_n) = \psi(s_{j1}, s_2, \dots, s_n)$$

for $j = 1, 2, \dots$. Since $\hat{\mu}(z, s_2, \dots, s_n)$ and $\psi(z, s_2, \dots, s_n)$ are analytic functions of z in the whole complex plane and they agree on the infinite bounded set $\{s_{j1}, j = 1, 2, \dots\}$ it follows that

$$\hat{\mu}(z, s_2, \dots, s_n) = \psi(z, s_2, \dots, s_n), \quad \forall z \in \mathbb{C}.$$

Since this holds for all $(s_2, \dots, s_n)^T \in \mathbb{D}$ which is dense in $\mathbb{R}^{(n-1)}$ and both sides of the equation are continuous on \mathbb{R}^n we have

$$\hat{\mu}(s_1, s_2, \dots, s_n) = \psi(s_1, s_2, \dots, s_n)$$

for all $(s_1, s_2, \dots, s_n)^T \in \mathbb{R}^n$. This implies that μ is a Gaussian measure with mean vector \mathbf{m} and covariance matrix Σ . \square

Acknowledgements

We thank S Ramasubramanian for bringing our attention to the example of E. Nelson on page 99 of [1] and J Stoyanov for useful suggestions.

References

- [1] Feller W, An Introduction to Probability Theory and Its Applications (2000) (India: Wiley) vol. II, 2nd edition
- [2] Hamedani G G and Tata M N, On the determination of the bivariate normal distribution from distributions of linear combinations of the variables, *The Amer. Math. Mon.* **82** (1975) 913–915
- [3] Kale B K, Normality of linear combinations of non-normal random variables, *The Amer. Math. Mon.* **77** (1970) 992–995
- [4] Shao Y and Zhou M, A characterization of multivariate normality through univariate projections, *J. Multi. Anal.* **101** (2010) 2637–2640
- [5] Stoyanov J, Counterexamples in Probability (1997) (New York: Wiley) 2nd edition