isid/ms/2011/16
September 6, 2011
http://www.isid.ac.in/~statmath/eprints

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

B. G. MANJUNATH
K. R. PARTHASARATHY

Indian Statistical Institute, Delhi Centre
7, SJSS Marg, New Delhi–110 016, India

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 – 110016, India
E–mail : bgmanjunath@gmail.com; krp@isid.ac.in

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.

Key words. gaussian distribution, characteristic function, homogeneous polynomial, linear functionals, nonunimodality, Hermite polynomial

AMS 1991 subject classification: primary, 60G15, 60E10; secondary, 62E15

1 Introduction

Starting with the simple example of E. Nelson as cited by W. Feller in [1] we have from the papers of B.K. Kale [3], G.G. Hamedani and M.N. Tata [2] and Y. Shao and M. Zhou [4] etc., as well as Section 10 of J. Stoyanov's book [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 standard gaussian in each S_j, $j \in \{1, 2, ..., N\}$; (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 coordinates has its distribution preserved. In particular, the sum of squares of all the coordinates can have the χ^2 distribution with n degrees of freedom.

We also demonstrate the following characterization of the multivariate gaussian distribution. Suppose $\{S_j, j = 1, 2, ...\}$ 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 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 $\boldsymbol{I}$ as follows. Choose and fix any homogeneous polynomial $\mathcal{P}$ of even degree $2k$ in n real variables $t_1, t_2, ..., t_n$ and define

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

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

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

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

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

Let

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

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 $: \mathcal{P} :$ 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

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

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

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

its renormalized version $:\mathcal{P}:$ is defined by

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

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

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

First, we note that

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

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

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

Thus the inverse Fourier transform (2.4) assumes the form

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

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

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

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

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

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 $\boldsymbol{x}\longmapsto\boldsymbol{a}^T\boldsymbol{x}$ with $|\boldsymbol{a}|=1$ has characteristic function $\varphi_{\boldsymbol{a}}$ and probability density function $g_{\boldsymbol{a}}$ on the real line given respectively by*

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

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

In particular, for any $\boldsymbol{a}\in Z_{\mathcal{P}}$, the linear functional $\boldsymbol{a}^T\boldsymbol{x}$ has the normal distribution with mean 0 and variance $|\boldsymbol{a}|^2$ 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 $\varphi_{\boldsymbol{a}}(t)$ of the linear functional $\boldsymbol{a}^T\boldsymbol{x}$ under the density function $f(\cdot;\varepsilon,\sigma,\mathcal{P})$ is $\Phi(t\boldsymbol{a};\varepsilon,\sigma,\mathcal{P})$ and (2.7) follows from (2.1) and the homogeneity of $\mathcal{P}$. Now (2.8) follows from Fourier inversion of (2.7). If $0\neq\boldsymbol{a}\in Z_{\mathcal{P}}$ then $0=\mathcal{P}(\boldsymbol{a})=\mathcal{P}\left(\frac{\boldsymbol{a}}{|\boldsymbol{a}|}\right)$ and therefore

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

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

Corollary 2.3. *Let $\{S_j, 1 \le j \le N\}$ 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 $j \in \{1, 2, ..., N\}$.*

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

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

Clearly,

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

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 $f(\cdot; \varepsilon, \sigma, \mathcal{P})$, $0 \ne \varepsilon$ 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 $\boldsymbol{b}^T \boldsymbol{x}$ has a normal distribution with mean 0 and variance $|\boldsymbol{b}|^2$ whenever $\boldsymbol{b} \in Z_{\mathcal{P}}$. This completes the proof. □

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

$$\left\{ x \middle| x \ne 0, g'_{\boldsymbol{a}}(x) = 0 \right\} =$$
$$\left\{ x \middle| x \ne 0, e^{\frac{x^2}{2}(\frac{1}{\sigma^2} - 1)} + \frac{(-1)^k \varepsilon \mathcal{P}(\boldsymbol{a})}{\sigma^{2k+2}} \frac{H_{2k+1}(\frac{x}{\sigma})}{x} = 0 \right\}.$$

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

Example 2.5. Let n be even,

$$\mathcal{P}(t_1,t_2,...,t_n) = t_1t_2...t_n\prod_{i>j}(t_i^2-t_j^2)$$

$$= t_1t_2...t_n\begin{vmatrix} 1 & 1 & ... & 1 \\ t_1^2 & t_2^2 & ... & t_n^2 \\ t_1^4 & t_2^4 & ... & t_n^4 \\ . & . & ... & . \\ . & . & ... & . \\ . & . & ... & . \\ t_1^{2(n-1)} & t_2^{2(n-1)} & ... & t_n^{2(n-1)} \end{vmatrix} \tag{2.9}$$

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

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

In particular, $:\mathcal{P}:$ is antisymmetric in the variables $x_1, x_2, ..., x_n$. 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 $f(\cdot;\varepsilon,\sigma,\mathcal{P})$ given by

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

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

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

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

Thus

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

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 $f(\boldsymbol{x}; \varepsilon, \sigma, \mathcal{P})$ in (2.11) is the same as its distribution under the standard gaussian density function with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{I}$.

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

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

This supremum is easily evaluated by switching over to the polar coordinates $x_1 = rcos\theta$, $x_2 = rsin\theta$. Then

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

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 $\boldsymbol{I}$. We write $\eta = 32 \; \varepsilon$ and express the density function (2.12) as

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

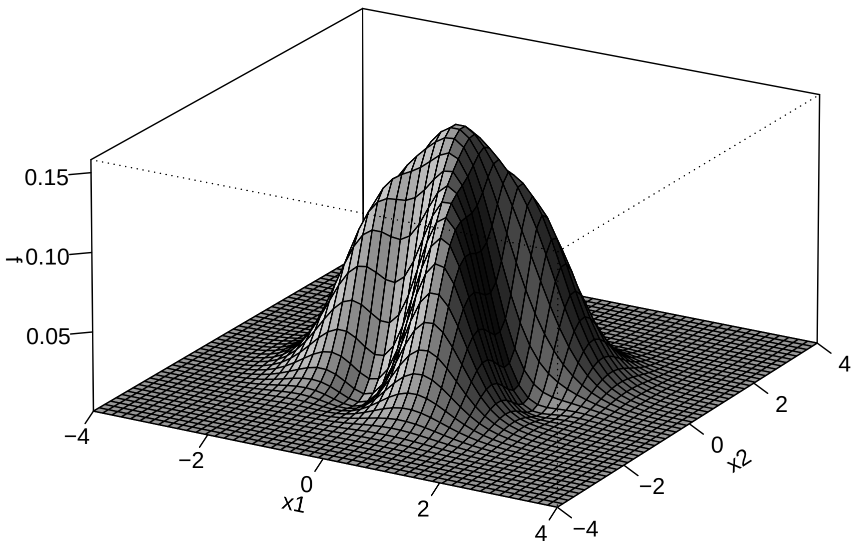

Figure 1: Bivariate density $f_\eta(x_1, x_2)$ at $\eta = e^2/4$.

where

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

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

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

Thus

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

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

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

This reveals the possibility of nonunimodality of the density of some linear functionals under the joint density f_η. For an illutration cf. Fig. (2).

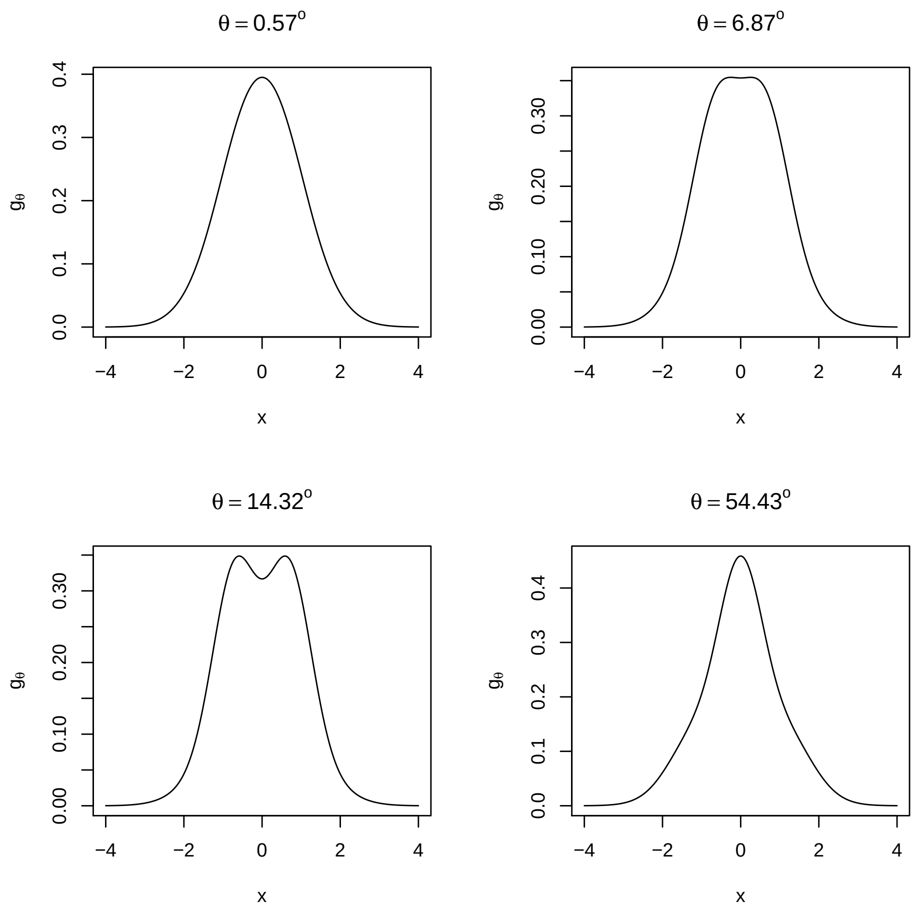

Figure 2: Nonunimodality of g_θ.

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 $\{S_j, j = 1, 2...\}$ 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 $j = 1, 2, \ldots$. 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 $\widehat{\mu}$ of μ given by

$$\widehat{\mu}(z_1, ..., z_n) = \int \exp(z_1 x_1 + ... + z_n x_n)\mu(dx_1 dx_2 ... dx_n) \tag{3.1}$$

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

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

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

Since

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

it follows that $\alpha_j \ge n^{-1/2}, \forall\, j$. 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

$$\begin{aligned} \boldsymbol{a}^{(j)} &= (a_{j1}, ..., a_{jn})^T, \\ |a_{j1}| &= \max_{1 \le r \le n} |a_{jr}|\ \forall\, j = 1, 2, ..., \\ |a_{j1}| &\ge n^{-1/2}\ \forall j. \end{aligned}$$

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

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

where

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

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

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

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

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

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

$$s_{j1} = -\boldsymbol{b}^{(j)^T} \boldsymbol{s},\ j = 1, 2, ...$$

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

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

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

$$\psi(z_1, ..., z_n) = \exp(\boldsymbol{m}^T\boldsymbol{z} + \frac{1}{2}\boldsymbol{z}^T\Sigma\boldsymbol{z}),\ \boldsymbol{z} \in \mathbb{C}^n$$

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

$$\widehat{\mu}(s_{j1}, s_2, ..., s_n) = \psi(s_{j1}, s_2, ..., s_n)$$

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

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

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

$$\widehat{\mu}(s_1, s_2, ..., s_n) = \psi(s_1, s_2, ..., s_n)$$

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

Acknowledgement

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

References

[1] W. Feller, An Introduction to Probability Theory and Its Applications (2000) (India: Wiley), Vol. II, 2nd edition

[2] G.G. Hamedani and M.N. Tata, On the determination of the bivariate normal distribution from distributions of linear combinations of the variables, The American Mathematical Monthly, 82 (1975), 913–915

[3] B.K. Kale, Normality of linear combinations of non-normal random variables, The American Mathematical Monthly, 77 (1970), 992–995

[4] Y. Shao and M. Zhou, A characterization of multivariate normality through univariate projections, J. Multi. Analysis, 101 (2010), 2637–2640

[5] J. Stoyanov, Counterexamples in Probability (1997) (New York: Wiley), 2nd edition