Lin. Algeb. and Appl. **321**(2001), 9–25

AN APPLICATION OF THE MATRIX VOLUME IN PROBABILITY

ADI BEN-ISRAEL

Dedicated to Professor Hans Schneider on his Seventieth Birthday

ABSTRACT. Given an n-dimensional random variable $\mathbf{X}$ with a joint density $f_{\mathbf{X}}(x_1, \cdots, x_n)$, the density of $\mathbf{Y} = h(\mathbf{X})$ is computed as a surface integral of $f_{\mathbf{X}}$ in two cases: (a) h linear, and (b) h sum of squares. The integrals use the volume of the Jacobian matrix in a change-of-variables formula.

1. INTRODUCTION

The abbreviation RV of *Random Variable* is used throughout. Consider n RV's $(\mathbf{X}_1, \cdots, \mathbf{X}_n)$ with a given joint density $f_{\mathbf{X}}(x_1, \cdots, x_n)$ and a RV

$$\mathbf{Y} = h(\mathbf{X}_1, \cdots, \mathbf{X}_n) \tag{1}$$

defined by the mapping $h : \mathbb{R}^n \to \mathbb{R}$. The density function $f_{\mathbf{Y}}(y)$ of $\mathbf{Y}$ is derived here in two special cases,

$$h \text{ linear}: \ h(\mathbf{X}_1, \cdots, \mathbf{X}_n) = \sum_{i=1}^{n} \xi_i \mathbf{X}_i \ , \text{ see Corollary 1} \ , \tag{2}$$

$$h \text{ sum of squares}: \ h(\mathbf{X}_1, \cdots, \mathbf{X}_n) = \sum_{i=1}^{n} \mathbf{X}_i^2 \ , \text{ Corollary 2} \ . \tag{3}$$

In both cases, the density $f_{\mathbf{Y}}(y)$ is computed as an integral of $f_{\mathbf{X}}$ on the surface

$$\mathcal{V}(y) := \{\mathbf{x} \in \mathbb{R}^n : h(\mathbf{x}) = y\} \ ,$$

that is a hyperplane for (2) and a sphere for (3). These integrals are elementary, and computationally feasible, as illustrated in Appendix A with the computer algebra package DERIVE [5]. Both results are consequences of Theorem 1, and a comparison between two integrations, one "classical" and the other using the change-of-variables formula of [2],

$$\int_{\mathcal{V}} f(\mathbf{v})\, d\mathbf{v} = \int_{\mathcal{U}} (f \circ \phi)(\mathbf{u}) \operatorname{vol} J_\phi(\mathbf{u})\, d\mathbf{u} \ , \tag{4}$$

where:

- $\mathcal{U} \subset \mathbb{R}^n$, $\mathcal{V} \subset \mathbb{R}^m$ and $m \geq n$,
- f a real-valued function integrable on $\mathcal{V}$,
- ϕ is a sufficiently well-behaved function: $\mathcal{U} \to \mathcal{V}$,
- $\circ$ denotes *composition*, here $(f \circ \phi)(\mathbf{u}) := f(\phi(\mathbf{u}))$,

Date: 13 June 1998. Revised May 30, 2000.

1991 *Mathematics Subject Classification.* Primary 15A15, 60D05; Secondary 26B15.

Key words and phrases. Determinants, Jacobians, matrix volume, change-of-variables in integration, surface integrals, geometric probabilities, distributions, Radon transforms.

- J_ϕ is the *Jacobi matrix* (or *Jacobian*) of ϕ,
$$J_\phi := \left(\frac{\partial \phi_i}{\partial u_j}\right) , \quad \text{also denoted} \quad \frac{\partial(v_1, v_2, \cdots, v_m)}{\partial(u_1, u_2, \cdots, u_n)} ,$$
representing the derivative of ϕ,
- J_ϕ is assumed of full-column rank throughout $\mathcal{U}$, and
- $\operatorname{vol} J_\phi$ denotes the *volume* of J_ϕ, see e.g. [1].

Recall that the volume of an $m \times n$ matrix of rank r is
$$\operatorname{vol} A := \sqrt{\sum_{(I,J)\in\mathcal{N}} \det^2 A_{IJ}} , \tag{5}$$
where A_{IJ} is the submatrix of A with rows I and columns J, and $\mathcal{N}$ is the index set of $r \times r$ nonsingular submatrices of A. Alternatively, $\operatorname{vol} A$ is the product of the singular values of A. If A is of full column rank, its volume is simply
$$\operatorname{vol} A = \sqrt{\det A^T A} . \tag{6}$$
If $m = n$ then $\operatorname{vol} J_\phi = |\det J_\phi|$, and (4) gives the classical result,
$$\int_{\mathcal{V}} f(\mathbf{v})\, d\mathbf{v} \;=\; \int_{\mathcal{U}} (f \circ \phi)(\mathbf{u})\; |\det J_\phi(\mathbf{u})|\; d\mathbf{u} . \tag{7}$$
The change-of-variables formula (4) is used here for surface integrals, over surfaces $\mathcal{S}$ in $\mathbb{R}^n$ given by
$$x_n \;:=\; g(x_1, x_2, \ldots, x_{n-1}) . \tag{8}$$
Let $\mathcal{V}$ be a subset on $\mathcal{S}$, and let $\mathcal{U}$ be the projection of $\mathcal{V}$ on $\mathbb{R}^{n-1}$, the space of variables $(x_1, \ldots, x_{n-1})$. The surface $\mathcal{S}$ is the graph of the mapping $\phi : \mathcal{U} \to \mathcal{V}$, given by its components $\phi \;:=\; (\phi_1, \phi_2, \ldots, \phi_n)$,
$$\begin{aligned} \phi_i(x_1, \ldots, x_{n-1}) \;&:=\; x_i \;, \quad i \in \overline{1, n-1} \\ \phi_n(x_1, \ldots, x_{n-1}) \;&:=\; g(x_1, \ldots, x_{n-1}) \end{aligned}$$
The Jacobi matrix of ϕ is
$$J_\phi \;=\; \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ g_{x_1} & g_{x_2} & \cdots & g_{x_{n-2}} & g_{x_{n-1}} \end{pmatrix}$$
where subscripts denote partial differentiation: $g_{x_1} = \frac{\partial g}{\partial x_1}$, etc. The volume is
$$\operatorname{vol} J_\phi \;=\; \sqrt{1 + \sum_{i=1}^{n-1} g_{x_i}^2} \tag{9}$$
For any function f integrable on $\mathcal{V}$ we therefore have, from (4),
$$\int\limits_{\mathcal{V}} f(x_1, \cdots, x_{n-1}, x_n)\, dV = \int\limits_{\mathcal{U}} f(x_1, \cdots, x_{n-1}, g(x_1, \cdots, x_{n-1})) \sqrt{1 + \sum_{i=1}^{n-1} g_{x_i}^2}\, dx_1 \cdots dx_{n-1} \tag{10}$$
Notation: We use the Euclidean norm $\|(x_1, \cdots, x_n)\| := \sqrt{\sum_{i=1}^n |x_i|^2}$.
For a random variable $\mathbf{X}$ we denote

$F_{\mathbf{X}}(x) := \text{Prob}\{\mathbf{X} \le x\}$, the distribution function, $f_{\mathbf{X}}(x) := \frac{d}{dx} F_{\mathbf{X}}(x)$, the density function,
$\text{E}\{\mathbf{X}\}$, the expected value, $\text{Var}\{\mathbf{X}\}$, the variance,
$\mathbf{X} \sim U(S)$ the fact that $\mathbf{X}$ is uniformly distributed over the set S,
$\mathbf{X} \sim N(\mu, \sigma)$ normally distributed with $\text{E}\{\mathbf{X}\} = \mu$, $\text{Var}\{\mathbf{X}\} = \sigma^2$,
$\mathbf{X} \sim \beta(p, q)$ beta(p,q) distributed, see (B.5).

Blanket assumption: Throughout this paper all random variables are absolutely continuous, and the indicated densities exist.

2. Probability densities and surface integrals

Let $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)$ have joint density $f_{\mathbf{X}}(x_1, \cdots, x_n)$ and let

$$y = h(x_1, \cdots, x_n) \tag{11}$$

where $h : \mathbb{R}^n \to \mathbb{R}$ is sufficiently well-behaved, in particular $\frac{\partial h}{\partial x_n} \ne 0$, and (11) can be solved for x_n,

$$x_n = h^{-1}(y | x_1, \cdots, x_{n-1}) \tag{12}$$

with $x_1, \cdots, x_{n-1}$ as parameters. By changing variables from $\{x_1, \cdots, x_n\}$ to $\{x_1, \cdots, x_{n-1}, y\}$, and using the fact

$$\det \left(\frac{\partial(x_1, \cdots, x_n)}{\partial(x_1, \cdots, x_{n-1}, y)} \right) = \frac{\partial h^{-1}}{\partial y} \tag{13}$$

we write the density of $\mathbf{Y} = h(\mathbf{X}_1, \cdots, \mathbf{X}_n)$ as

$$f_{\mathbf{Y}}(y) = \int\limits_{\mathbb{R}^{n-1}} f_{\mathbf{X}}(x_1, \cdots, x_{n-1}, h^{-1}(y | x_1, \cdots, x_{n-1})) \left| \frac{\partial h^{-1}}{\partial y} \right| dx_1 \cdots dx_{n-1} \tag{14}$$

Let $\mathcal{V}(y)$ be the surface given by (11), represented as

$$\begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ h^{-1}(y | x_1, \cdots, x_{n-1}) \end{pmatrix} = \phi \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \end{pmatrix} \tag{15}$$

Then the surface integral of $f_{\mathbf{X}}$ over $\mathcal{V}(y)$ is given, by (10), as

$$\int\limits_{\mathcal{V}(y)} f_{\mathbf{X}} = \int\limits_{\mathbb{R}^{n-1}} f_{\mathbf{X}}(x_1, \cdots, x_{n-1}, h^{-1}(y | \, x_1, \cdots, x_{n-1})) \sqrt{1 + \sum_{i=1}^{n-1} \left(\frac{\partial h^{-1}}{\partial x_i} \right)^2} \, dx_1 \cdots dx_{n-1} \tag{16}$$

Theorem 1. If the ratio

$$\frac{\frac{\partial h^{-1}}{\partial y}}{\sqrt{1 + \sum\limits_{i=1}^{n-1} \left(\frac{\partial h^{-1}}{\partial x_i} \right)^2}} \quad \text{does not depend on } x_1, \cdots, x_{n-1} \,, \tag{17}$$

then

$$f_{\mathbf{Y}}(y) = \frac{\left| \frac{\partial h^{-1}}{\partial y} \right|}{\sqrt{1 + \sum\limits_{i=1}^{n-1} \left(\frac{\partial h^{-1}}{\partial x_i} \right)^2}} \int\limits_{\mathcal{V}(y)} f_{\mathbf{X}} \tag{18}$$

Proof. A comparison of (16) and (14) gives the density $f_{\mathbf{Y}}$ as the surface integral (18), $\square$

Condition (17) holds if $\mathcal{V}(y)$ is a hyperplane (see § 3) or a sphere, see § 4. In these two cases, covering many important probability distributions, the derivation (18) is simpler computationally than classical integration formulae, e.g. [3, Theorem 5.1.5], [7, Theorem 6-5.4] and transform methods, e.g. [10].

3. Hyperplanes

Let

$$y = h(x_1, \cdots, x_n) := \sum_{i=1}^{n} \xi_i \, x_i \tag{19}$$

where $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_n)$ is a given vector with $\xi_n \neq 0$. Then (12) becomes

$$x_n = h^{-1}(y | \, x_1, \cdots, x_{n-1}) \;\; := \;\; \frac{y}{\xi_n} - \sum_{i=1}^{n-1} \frac{\xi_i}{\xi_n} \, x_i \tag{20a}$$

$$\text{with} \quad \frac{\partial h^{-1}}{\partial y} = \frac{1}{\xi_n} \, , \quad \sqrt{1 + \sum_{i=1}^{n-1} \left(\frac{\partial h^{-1}}{\partial x_i} \right)^2} \;\; = \;\; \sqrt{1 + \sum_{i=1}^{n-1} \left(\frac{\xi_i}{\xi_n} \right)^2} = \frac{\|\boldsymbol{\xi}\|}{|\xi_n|} \tag{20b}$$

Condition (17) thus holds, and the density of $\sum \xi_i \, \mathbf{X}_i$ can be expressed as a surface integral of $f_{\mathbf{X}}$ on the hyperplane

$$\mathcal{H}(\boldsymbol{\xi}, y) := \left\{ \mathbf{x} \in \mathbb{R}^n : \sum_{i=1}^{n} \xi_i \, x_i \; = \; y \right\} \tag{21}$$

also called the *Radon transform* $(\mathbf{R} f_{\mathbf{X}})(\boldsymbol{\xi}, y)$ of $f_{\mathbf{X}}$. The Radon transform can be computed as an integral on $\mathbb{R}^{n-1}$, see [2, Example 7],

$$(\mathbf{R} f_{\mathbf{X}})(\boldsymbol{\xi}, y) \; = \; \frac{\|\boldsymbol{\xi}\|}{|\xi_n|} \int_{\mathbb{R}^{n-1}} f_{\mathbf{X}} \left(x_1, \cdots, x_{n-1}, \frac{y}{\xi_n} - \sum_{i=1}^{n-1} \frac{\xi_i}{\xi_n} \, x_i \right) dx_1 dx_2 \cdots dx_{n-1} \, . \tag{22}$$

The cases $n = 2$ and $n = 3$ are computed in (A.1) and (A.2) below.
Note: The normal vector $\boldsymbol{\xi}$ of the hyperplane (21) can be normalized, and can therefore be assumed a unit vector, see e.g. [4, Chapter 3] where the Radon transform with respect to a hyperplane

$$\mathcal{H}(\boldsymbol{\xi}^0, p) := \{ \mathbf{x} \in \mathbb{R}^n : \, \boldsymbol{\xi}^0 \cdot \mathbf{x} = p \} \; , \; \|\boldsymbol{\xi}^0\| = 1 \, , \tag{23}$$

is represented as

$$\check{f}(p, \boldsymbol{\xi}^0) = \int f(\mathbf{x}) \, \delta(p - \boldsymbol{\xi}^0 \cdot \mathbf{x}) \, d\mathbf{x} \, ,$$

where $\delta(\cdot)$ is the Dirac delta function. If (21) and (23) represent the same hyperplane, then the correspondence between $(\mathbf{R} f_{\mathbf{X}})(\boldsymbol{\xi}, y)$ and $\check{f}(p, \boldsymbol{\xi}^0)$ is given by

$$\boldsymbol{\xi}^0 = \frac{\boldsymbol{\xi}}{\|\boldsymbol{\xi}\|} \; , \; p = \frac{y}{\|\boldsymbol{\xi}\|} \, .$$

Corollary 1. Let $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n)$ be random variables with joint density $f_{\mathbf{X}}(x_1, x_2, \cdots, x_n)$, and let $\mathbf{0} \neq \boldsymbol{\xi} \in \mathbb{R}^n$. The random variable

$$\mathbf{Y} \; := \; \sum_{i=1}^{n} \xi_i \, \mathbf{X}_i \tag{24}$$

has the density

$$f_{\mathbf{Y}}(y) \; = \; \frac{(\mathbf{R} f_{\mathbf{X}})(\boldsymbol{\xi}, y)}{\|\boldsymbol{\xi}\|} \, . \tag{25}$$

Proof. Follows from (18), (20b) and (22). □

Explanation of the factor $\|\boldsymbol{\xi}\|$ in (25): the distance between the hyperplanes $\mathcal{H}(\boldsymbol{\xi}, y)$ and $\mathcal{H}(\boldsymbol{\xi}, y + dy)$ is $dy/\|\boldsymbol{\xi}\|$.

Example 1 (*Bivariate normal distribution*). Let $(\mathbf{X}_1, \mathbf{X}_2)$ have the bivariate normal distribution with zero means and unit variances,

$$f_{\mathbf{X}}(x_1, x_2) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{x_1^2 - 2\rho x_1 x_2 + x_2^2}{2(1-\rho^2)}\right\} \tag{26}$$

and let $\mathbf{Y} := a\,\mathbf{X}_1 + b\,\mathbf{X}_2$. The density of $\mathbf{Y}$ is, by (25),

$$\begin{aligned} f_{\mathbf{Y}}(y) &= \frac{1}{\sqrt{a^2+b^2}}\,(\mathbf{R} f_{\mathbf{X}})((a,b), y) \\ &= \frac{1}{\sqrt{2\,\pi}\,\sqrt{a^2 + 2\,a\,b\,\rho + b^2}} \exp\left\{-\frac{y^2}{2\,(a^2 + 2\,a\,b\,\rho + b^2)}\right\}\ , \quad \text{see (A.5)}\ . \end{aligned} \tag{27}$$

Therefore $a\,\mathbf{X}_1 + b\,\mathbf{X}_2 \ \sim\ N(0, \sqrt{a^2 + 2\,a\,b\,\rho + b^2})$. In particular, $\mathbf{X}_1 + \mathbf{X}_2 \sim N(0, \sqrt{2(1+\rho)})$ and $\mathbf{X}_1 - \mathbf{X}_2 \sim N(0, \sqrt{2(1-\rho)})$.

Example 2 (*Uniform distribution*). Let $(\mathbf{X}_1, \mathbf{X}_2)$ be independent and uniformly distributed on $[0,1]$. Their joint density is

$$f_{\mathbf{X}}(x_1, x_2) = \begin{cases} 1 & ,\ \text{if } 0 \le x_1, x_2 \le 1 \\ 0 & ,\ \text{otherwise} \end{cases}$$

and the density of $a\,\mathbf{X}_1 + b\,\mathbf{X}_2$ is, by (25),

$$\begin{aligned} f_{a\,\mathbf{X}_1 + b\,\mathbf{X}_2}(y) &= \frac{1}{\sqrt{a^2+b^2}}\,(\mathbf{R} f_{\mathbf{X}})((a,b), y) \\ &= \frac{|y-a-b| - |y-a| - |y-b| + |y|}{2\,a\,b}\ , \quad \text{see (A.7)}\ . \end{aligned} \tag{28}$$

In particular,

$$f_{\mathbf{X}_1+\mathbf{X}_2}(y) = \begin{cases} y & ,\ \text{if } 0 \le y < 1 \\ 2-y & ,\ \text{if } 1 \le y \le 2 \\ 0 & ,\ \text{otherwise} \end{cases}$$

a symmetric triangular distribution on $[0,2]$.

4. Spheres

Let

$$\begin{aligned} \mathcal{B}_n(r) &:= \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \le r\}\ , \quad \text{the } \textit{ball} \text{ of radius } r, \\ \mathcal{S}_n(r) &:= \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = r\}\ , \quad \text{the } \textit{sphere} \text{ of radius } r, \end{aligned}$$

both centered at the origin. Also,

$$\begin{aligned} v_n(r) &:= \text{the } \textit{volume} \text{ of } \mathcal{B}_n(r)\ , \\ a_n(r) &:= \text{the } \textit{area} \text{ of } \mathcal{S}_n(r)\ , \end{aligned}$$

where r is dropped if $r = 1$, so that

$$\begin{aligned} v_n &:= \text{the volume of the unit ball } \mathcal{B}_n\ , \\ a_n &:= \text{the area of the unit sphere } \mathcal{S}_n\ . \end{aligned}$$

Clearly

$$v_n(r) = v_n\,r^n\ , \quad a_n(r) = a_n\,r^{n-1}\ , \quad \text{and} \quad dv_n(r) = v_n'(r)\,dr = a_n(r)\,dr\ , \tag{29}$$

and it follows that

$$a_n \ =\ n\,v_n\ , \quad n = 2, 3, \cdots \tag{30}$$

The area of the unit sphere $\mathcal{S}_n$ is computed in Example 7 as

$$a_n = \frac{2\,\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)}\ , \quad n = 2, 3, \cdots \tag{31}$$

Let

$$y = h(x_1, \cdots, x_n) := \sum_{i=1}^{n} x_i^2 \tag{32}$$

which has two solutions (12) for x_n, representing the upper and lower hemispheres,

$$x_n = h^{-1}(y|\, x_1, \cdots, x_{n-1}) \ := \ \pm\sqrt{y - \sum_{i=1}^{n-1} x_i^2} \tag{33a}$$

$$\text{with} \quad \frac{\partial h^{-1}}{\partial y} = \pm\frac{1}{2\sqrt{y - \sum_{i=1}^{n-1} x_i^2}}\ , \quad \sqrt{1 + \sum_{i=1}^{n-1}\left(\frac{\partial h^{-1}}{\partial x_i}\right)^2} = \frac{\sqrt{y}}{\sqrt{y - \sum_{i=1}^{n-1} x_i^2}} \tag{33b}$$

Therefore condition (17) holds, and the density of $\sum \mathbf{X}_i^2$ is, by (18), expressed in terms of the surface integral of $f_{\mathbf{X}}$ on the sphere $\mathcal{S}_n(\sqrt{y})$ of radius $\sqrt{y}$.

Corollary 2. Let $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)$ have joint density $f_{\mathbf{X}}(x_1, \cdots, x_n)$. The density of

$$Y = \sum_{i=1}^{n} \mathbf{X}_i^2 \tag{34}$$

is

$$f_{\mathbf{Y}}(y) = \frac{1}{2\sqrt{y}} \int\limits_{\mathcal{S}_n(\sqrt{y})} f_{\mathbf{X}} \tag{35}$$

where the integral is over the sphere $\mathcal{S}_n(\sqrt{y})$ of radius $\sqrt{y}$, computed as an integral over the ball $\mathcal{B}_{n-1}(\sqrt{y})$,

$$\int\limits_{\mathcal{S}_n(\sqrt{y})} f_{\mathbf{X}} =$$

$$\int\limits_{\mathcal{B}_{n-1}(\sqrt{y})} \left[f_{\mathbf{X}}\left(x_1, \cdots, x_{n-1}, \sqrt{y - \sum_{i=1}^{n-1} x_i^2}\right) + f_{\mathbf{X}}\left(x_1, \cdots, x_{n-1}, -\sqrt{y - \sum_{i=1}^{n-1} x_i^2}\right)\right] \frac{\sqrt{y}\ dx_1 \cdots dx_{n-1}}{\sqrt{y - \sum_{i=1}^{n-1} x_i^2}} \tag{36}$$

Proof. (35) follows from (18) and (33b). The surface integral (36) is (10) with $g = h^{-1}$. □

An explanation of the factor $2\sqrt{y}$ in (35): the width of the spherical shell bounded by the two spheres $\mathcal{S}_n(\sqrt{y})$ and $\mathcal{S}_n(\sqrt{y + dy})$ is the difference of radii

$$\sqrt{y + dy} - \sqrt{y} \approx \frac{dy}{2\sqrt{y}}$$

Example 3 (*Spherical distribution*)**.** If the joint density of $\mathbf{X} = (\mathbf{X}_1, \cdots, \mathbf{X}_n)$ is spherical

$$f_{\mathbf{X}}(x_1, \cdots, x_n) = p\left(\sum_{i=1}^{n} x_i^2\right) \tag{37}$$

then $\mathbf{Y} = \sum_{i=1}^{n} \mathbf{X}_i^2$ has the density

$$f_{\mathbf{Y}}(y) = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2})}\, p(y)\, y^{\frac{n}{2}-1} \tag{38}$$

Proof. The surface integral of $f_{\mathbf{X}}$ over $\mathcal{S}_n(\sqrt{y})$ is

$$\int\limits_{\mathcal{S}_n(\sqrt{y})} f_{\mathbf{X}} = p(y)\, a_{n-1}(\sqrt{y})$$

$$= p(y)\, \frac{2\,\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)}\, \sqrt{y}^{\,n-1}$$

by (31) and (29). The proof is completed by (35). □

Example 4 (χ^2 *distribution*)**.** If $\mathbf{X}_i \sim N(0,1)$ and are independent, $i \in \overline{1,n}$, their joint density is of the form (37),

$$f_{\mathbf{X}}(x_1,\cdots,x_n) = (2\pi)^{-n/2}\ \exp\left\{-\frac{\sum_{i=1}^n x_i^2}{2}\right\}$$

and (38) gives,

$$f_{\mathbf{Y}}(y) = \frac{1}{2^{n/2}\,\Gamma(\frac{n}{2})}\, y^{n/2-1}\ \exp\left\{-\frac{y}{2}\right\}\ , \tag{39}$$

the χ^2 *distribution* with n degrees of freedom.

Example 5 (*Random directions in* $\mathbb{R}^n$, [6, § I.10])**.** If $\mathbf{X} \sim U(\mathcal{S}_n)$ (uniform distribution on the unit sphere) the probability of a surface element on $\mathcal{S}_n$ is, by (C.1),

$$\frac{dx_1 dx_2 \cdots dx_{n-1}}{a_n \sqrt{1-\sum_{i=1}^{n-1} x_i^2}}\ . \tag{40}$$

Random points on $\mathcal{S}_n$ are also called *random directions*. We study the distributions of projections of random directions on lines and hyperplanes.
(a) Let $\mathbf{L}_n$ be the length of the projection of a unit vector in $\mathbb{R}^n$ on a fixed line through the origin, say the x_n–axis. Then $\mathbf{L}_n$ has the density

$$f_{\mathbf{L}_n}(x) = \frac{2}{B(\frac{1}{2},\frac{n-1}{2})}\,(1-x^2)^{\frac{n-3}{2}}\ , \tag{41a}$$

$$\text{and expected value}\quad \mathrm{E}\,\{\mathbf{L}_n\} = \frac{\Gamma(\frac{n}{2})}{\Gamma(\frac{1}{2})\,\Gamma(\frac{n+1}{2})}\ . \tag{41b}$$

The density (41a) is simplest when $n = 3$, in which case $\mathbf{L}_3 \sim U([0,1])$. This means that distances to a fixed plane through the center are uniformly distributed on $\mathcal{S}_3$.
(b) The square $\mathbf{L}_n^2$ has the beta distribution, see (B.5),

$$\mathbf{L}_n^2 \sim \beta\left(\frac{n-1}{2},\frac{1}{2}\right)\ , \tag{42a}$$

$$\text{with}\quad \mathrm{E}\,\{\mathbf{L}_n^2\} = \frac{1}{n}\quad ,\quad \mathrm{Var}\,\{\mathbf{L}_n^2\} = \frac{2\,(n-1)}{n^2\,(n+2)}\ . \tag{42b}$$

(c) Let $\mathbf{H}_n$ be the length of the projection of a unit vector in $\mathbb{R}^n$ on a fixed hyperplane passing through the origin, say the hyperplane orthogonal to the x_n–axis. Then $\mathbf{H}_n$ has the density

$$f_{\mathbf{H}_n}(x) = \frac{2}{B(\frac{1}{2},\frac{n-1}{2})}\,\frac{x^{n-2}}{\sqrt{1-x^2}}\ , \tag{43a}$$

$$\text{and expected value}\quad \mathrm{E}\,\{\mathbf{H}_n\} = \frac{2}{n-1}\,\frac{\Gamma^2(\frac{n}{2})}{\Gamma^2(\frac{n-1}{2})}\ . \tag{43b}$$

(d) The square $\mathbf{H}_n^2$ has the beta distribution

$$\mathbf{H}_n^2 \sim \beta\left(\frac{1}{2}, \frac{n-1}{2}\right) , \tag{44a}$$

$$\text{with} \quad \mathrm{E}\{\mathbf{H}_n^2\} = \frac{n-1}{n} \quad , \quad \mathrm{Var}\{\mathbf{H}_n^2\} = \frac{2\,(n-1)}{n^2\,(n+2)} . \tag{44b}$$

Proof. (a) From $|x_n| = \sqrt{1 - \sum_{i=1}^{n-1} x_i^2}$ it follows that $|x_n| \le t$ is equivalent to $\sum_{i=1}^{n-1} x_i^2 \ge 1 - t^2$. Let

$$A(t) := \text{area}\,\{\mathbf{x} \in \mathcal{S}_n : |x_n| \le t\}$$

Then $\mathbf{L}_n$ has the distribution function

$$F_{\mathbf{L}_n}(x) = \text{Prob}\,\{\mathbf{L}_n \le x\} = \frac{A(x)}{a_n} .$$

The area $A(x)$ is computed by

$$\begin{aligned} A(x) &= 2 \int\limits_{1-x^2 \le \sum_{i=1}^{n-1} x_i^2 \le 1} \frac{dx_1 \cdots dx_{n-1}}{\sqrt{1 - \sum_{i=1}^{n-1} x_i^2}} \\ &= 2\,a_{n-1} \int_{\sqrt{1-x^2}}^{1} \frac{r^{n-2}\,dr}{\sqrt{1-r^2}} , \quad \text{as in the computation of (C.3)} . \end{aligned} \tag{45}$$

The density of $\mathbf{L}_n$ is

$$f_{\mathbf{L}_n}(x) = \frac{d}{dx} F_{\mathbf{L}_n}(x) = \frac{A'(x)}{a_n} = \frac{2\,a_{n-1}}{a_n} \left(-\frac{(1-x^2)^{\frac{n-2}{2}}}{x} \right) \left(-\frac{x}{\sqrt{1-x^2}} \right) = \frac{2\,a_{n-1}}{a_n} (1-x^2)^{\frac{n-3}{2}}$$

and (41a) follows from (31). The expected value (41b) is obtained by routine integration.
(b) The distribution function of $\mathbf{L}_n^2$ is

$$F_{\mathbf{L}_n^2}(x) = \text{Prob}\,\{\mathbf{L}_n \le \sqrt{x}\} = \frac{2\,a_{n-1}}{a_n} \int_0^{\sqrt{x}} (1-y^2)^{\frac{n-3}{2}}\,dy .$$

Differentiation gives (42a). The formulas (42b) then follow from (B.6) with $p = \frac{n-1}{2}$, $q = \frac{1}{2}$.
(c) Let $A(t) = \text{area}\,\{\mathbf{x} \in \mathcal{S}_n : \sum_{i=1}^{n-1} x_i^2 \le t^2\}$. Then

$$\begin{aligned} A(x) &= 2 \int\limits_{0 \le \sum_{i=1}^{n-1} x_i^2 \le x^2} \frac{dx_1 \cdots dx_{n-1}}{\sqrt{1 - \sum_{i=1}^{n-1} x_i^2}} \\ &= 2\,(n-1)\,v_{n-1} \int_0^x \frac{r^{n-2}}{\sqrt{1-r^2}}\,dr \end{aligned} \tag{46}$$

The distribution function of $\mathbf{H}_n$ is

$$F_{\mathbf{H}_n}(x) = \text{Prob}\,\{\mathbf{H}_n \le x\} = \frac{A(x)}{a_n}$$

which is differentiated to give the density

$$\begin{aligned} f_{\mathbf{H}_n}(x) &= \frac{A'(x)}{a_n} \\ &= \frac{2\,a_{n-1}}{a_n} \frac{x^{n-2}}{\sqrt{1-x^2}} , \quad \text{by (46)} , \end{aligned}$$

proving (43a).
(d) The distribution function of $\mathbf{H}_n^2$ is

$$F_{\mathbf{H}_n^2}(x) = \text{Prob}\,\{\mathbf{H}_n \le \sqrt{x}\} = \frac{2\,a_{n-1}}{a_n} \int_0^{\sqrt{x}} \frac{y^{n-2}}{\sqrt{1-y^2}}\,dy\ .$$

Differentiation gives (44a). The formulas (44b) then follow from (B.6) with $p = \frac{1}{2}$, $q = \frac{n-1}{2}$. □

Combining (42b) and (44b) we conclude

$$\text{E}\,\{\mathbf{L}_n^2\} + \text{E}\,\{\mathbf{H}_n^2\} = 1$$

as expected, since for any unit vector $\mathbf{x} = (x_1, \cdots, x_{n-1})$,

$$x_n^2 + \left(\sum_{i=1}^{n-1} x_i^2 \right) = \mathbf{L}_n^2 + \mathbf{H}_n^2 = 1$$

This also explains why $\mathbf{L}_n^2$ and $\mathbf{H}_n^2$ have the same variance.

Example 6 (*Probabilistic proof*)**.** A probabilistic proof of (30) is given by writing

$$\begin{aligned} \frac{n\,v_n - a_n}{2\,a_n} &= \int\limits_{\mathcal{B}_{n-1}} \left[n \sqrt{1 - \sum_{k=1}^{n-1} x_k^2} \ - \ \frac{1}{\sqrt{1 - \sum_{k=1}^{n-1} x_k^2}} \right] \frac{dx_1 \cdots dx_{n-1}}{a_n}\ , \quad \text{by (C.5) and (C.2)}\ , \\ &= \int\limits_{\mathcal{B}_{n-1}} \left[(n-1) \ - \ n \sum_{k=1}^{n-1} x_k^2 \right] \frac{dx_1 \cdots dx_{n-1}}{a_n \sqrt{1 - \sum_{k=1}^{n-1} x_k^2}}\ . \end{aligned} \tag{47}$$

By (40), this is the expected value of $(n-1) - n \sum_{k=1}^{n-1} \mathbf{X}_k^2$, for $(\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n)$ uniformly distributed on $\mathcal{S}_n$. It then follows from (44b) that the RHS of (47) is zero, proving that $a_n = n\,v_n$.

References

[1] A. Ben-Israel, *A volume associated with $m \times n$ matrices*, Lin. Algeb. Appl. **167**(1992), 87–111.
[2] A. Ben-Israel, *The change of variables formula using matrix volume*, SIAM J. Matrix Analysis **21** (1999), 300–312.
[3] Z.W. Birnbaum, *Introduction to Probability and Mathematical Statistics*, Harper, 1962
[4] S.R. Deans, *The Radon Transform and Some of its Applications*, Wiley, 1983/Krieger 1993.
[5] Derive, Texas Instruments, http://www.ti.com/calc/docs/derive.htm
[6] W. Feller, *An Introduction to Probability Theory and its Applications*, Vol. 2, Wiley, 1966
[7] B. Harris, *Theory of Probability*, Addison-Wesley, 1966.
[8] C. Müller, *Spherical Harmonics*, Lecture Notes in Mathematics No. 17, Springer-Verlag, 1966.
[9] F. Natterer, *The Mathematics of Computerized Tomography*, J. Wiley, 1986.
[10] M.D. Springer, *The Algebra of Random Variables*, Wiley, 1979.

Appendix A: Illustrations with Derive.

The integrations of this paper can be done symbolically. We illustrate this for the symbolic package Derive, [5], omitting details such as the commands (e.g. `Simplify`, `Approximate`) and settings (e.g. `CaseMode:=Sensitive`, `InputMode:=Word`) that are used to obtain these results.
For convenience we define

$$\texttt{EVAL}(f, x, x_0) := \texttt{LIM}(f, x, x_0)$$

evaluating a function $f(x)$ at $x = x_0$.
(a) The Radon transform $(\mathbf{R}f)(\boldsymbol{\xi}, y)$ of (22) is computed for $n = 2$ by,

$$\texttt{RADON\_2}(f, x_1, x_2, a, b, y) := \int_{-\infty}^{\infty} \texttt{EVAL}\left(f,\ [x_1, x_2],\ \left[x_1, \frac{y}{b} - \frac{a}{b}\, x_1\right]\right)\ dx_1\ \frac{\sqrt{a^2+b^2}}{|b|} \tag{A.1}$$

where $\boldsymbol{\xi} = (a, b)$, and $b \neq 0$ is assumed.
The 3-dimensional Radon transform $(\mathbf{R}f)(\boldsymbol{\xi}, y)$ is computed by,

$$\texttt{RADON\_3}(f,\ x_1, x_2, x_3,\ a, b, c,\ y) :=$$
$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \texttt{EVAL}\left(f,\ [x_1, x_2, x_3],\ \left[x_1, x_2, \frac{y}{c} - \frac{a}{c}\, x_1 - \frac{b}{c}\, x_2\right]\right)\ dx_1\, dx_2\ \frac{\sqrt{a^2+b^2+c^2}}{|c|} \tag{A.2}$$

where $\boldsymbol{\xi} = (a, b, c)$, and $c \neq 0$.
(b) The density of a linear function $\sum \xi_i\, \mathbf{X}_i$ is, by (25),

$$\texttt{DENSITY\_2}(f, x_1, x_2, a, b, y) \ := \ \frac{\texttt{RADON\_2}(f, x_1, x_2, a, b, y)}{\sqrt{a^2+b^2}} \tag{A.3a}$$

$$\texttt{DENSITY\_3}(f, x_1, x_2, x_3, a, b, c, y) \ := \ \frac{\texttt{RADON\_3}(f, x_1, x_2, x_3, a, b, c, y)}{\sqrt{a^2+b^2+c^2}} \tag{A.3b}$$

for $n = 2$ and $n = 3$, respectively.
(c) The density in Example 1 is computed by,

$$\texttt{DENSITY\_2}\left(\frac{\texttt{EXP}\ \left[-\left(x_1^2 - 2\rho x_1 x_2 + x_2^2\right)/(2(1-\rho^2))\right]}{2\pi\sqrt{1-\rho^2}},\ x_1, x_2,\ a, b,\ y\right) \tag{A.4}$$

Declaring $\rho \in [-1, 1]$, (A.4) simplifies to

$$\frac{\sqrt{2}\,\texttt{EXP}\ \left[-\dfrac{y^2}{2\,(a^2 + 2\,a\,b\,\rho + b^2)}\right]}{2\,\sqrt{\pi}\,\sqrt{a^2 + 2\,a\,b\,\rho + b^2}} \tag{A.5}$$

proving (27).
(d) If $(\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3)$ are independent and $\sim N(0, 1)$ then the density of $a\,\mathbf{X}_1 + b\,\mathbf{X}_2 + c\,\mathbf{X}_3$ is computed by

$$\texttt{DENSITY\_3}\ \left((2\,\pi)^{-3/2}\,\texttt{EXP}\left[-\frac{x_1^2+x_2^2+x_3^2}{2}\right], x_1, x_2, x_3, a, b, c, y\right)$$

giving

$$\frac{\sqrt{2}\,\texttt{EXP}\ \left[-\frac{y^2}{2\,(a^2+b^2+c^2)}\right]}{2\,\sqrt{\pi}\,\sqrt{a^2+b^2+c^2}}$$

showing that $a\,\mathbf{X}_1 + b\,\mathbf{X}_2 + c\,\mathbf{X}_3 \ \sim\ N(0, \sqrt{a^2+b^2+c^2})$.
(e) In Example 2, the density of $a\,\mathbf{X}_1 + b\,\mathbf{X}_2$ is computed by

$$\texttt{DENSITY\_2}\,(\texttt{CHI}(0, x_1, 1)\texttt{CHI}(0, x_2, 1),\ x_1, x_2,\ a, b,\ y) \tag{A.6}$$

which gives

$$\frac{|y-a-b|}{2ab} - \frac{|y-a| + |y-b| - |y|}{2ab} \tag{A.7}$$

a rearrangement of (28). Similarly, if $\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3$ are independent and $\sim U([0, 1])$, the density of $a\,\mathbf{X}_1 + b\,\mathbf{X}_2 + c\,\mathbf{X}_3$ is computed by,

$$\texttt{DENSITY\_3}\,(\texttt{CHI}(0, x_1, 1)\texttt{CHI}(0, x_2, 1)\texttt{CHI}(0, x_3, 1),\ x_1, x_2, x_3,\ a, b, c,\ y)$$

giving

$$-\frac{(y-a-b-c)\,|y-a-b-c|}{4abc}+\frac{(y-a-b)\,|y-a-b|}{4abc}+\frac{(y-a-c)\,|y-a-c|}{4abc}$$
$$-\frac{(y-a)\,|y-a|-(y-b-c)\,|y-b-c|+(y-b)\,|y-b|+(y-c)\,|y-c|-y\,|y|}{4abc}$$

Appendix B: The Gamma and Beta Functions

The **gamma function** $\Gamma(p)$ is

$$\Gamma(p) := \int_0^\infty x^{p-1}\, e^{-x}\, dx\ . \tag{B.1}$$

Its properties include:

$$\Gamma(1) = 1 \tag{B.2a}$$
$$\Gamma(p+1) = p\,\Gamma(p) \tag{B.2b}$$
$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi} \tag{B.2c}$$

The **beta function** is

$$B(p,q) := \int_0^1 (1-x)^{p-1}\, x^{q-1}\, dx\ . \tag{B.3}$$

It satisfies:

$$B(p,q) = \frac{\Gamma(p)\,\Gamma(q)}{\Gamma(p+q)} \tag{B.4a}$$
$$\frac{B(p,q+1)}{B(p,q)} = \frac{q}{p+q} \tag{B.4b}$$
$$\frac{B(p+1,q)}{B(p,q)} = \frac{p}{p+q} \tag{B.4c}$$

where (B.4b)–(B.4c) follow from (B.4a) and (B.2b). A random variable $\mathbf{X}$ has the **beta**(p,q) distribution, denoted $\mathbf{X} \sim \beta(p,q)$, if its density is

$$\beta(x|\,p,q) := \frac{(1-x)^{p-1}\, x^{q-1}}{B(p,q)}\ , \quad 0 \le x \le 1\ , \tag{B.5}$$

in which case we verify, by repeat applications of (B.4b)–(B.4c),

$$\mathrm{E}\,\{\mathbf{X}\} = \frac{q}{p+q}\ , \quad \mathrm{Var}\,\{\mathbf{X}\} = \frac{p\,q}{(p+q)^2\,(p+q+1)}\ . \tag{B.6}$$

Note that $\beta(p,q) \neq \beta(q,p)$ if $p \neq q$, while $B(p,q) = B(q,p)$ for all p,q.

Appendix C: Spheres and balls in $\mathbb{R}^n$

Integrals on $\mathcal{S}_n$, in particular the area a_n, can be computed using **spherical coordinates**, e.g. [9, § VII.2], or the **surface element** of $\mathcal{S}_n$, e.g. [8]. An alternative, simpler, approach is to use (4), representing the "upper hemisphere" as $\phi(\mathcal{B}_{n-1})$, where $\phi = (\phi_1, \phi_2, \cdots, \phi_n)$ is

$$\phi_i(x_1, x_2, \cdots, x_{n-1}) = x_i\ ,\ i \in \overline{1, n-1}\ ,$$
$$\phi_n(x_1, x_2, \cdots, x_{n-1}) = \sqrt{1 - \sum_{i=1}^{n-1} x_i^2}\ .$$

The Jacobi matrix is

$$J_\phi = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ -\dfrac{x_1}{x_n} & -\dfrac{x_2}{x_n} & \cdots & -\dfrac{x_{n-1}}{x_n} \end{pmatrix}$$

and its volume is easily computed

$$\operatorname{vol} J_\phi = \sqrt{1 + \sum_{i=1}^{n-1} \left(\frac{x_i}{x_n}\right)^2} = \frac{1}{|x_n|} = \frac{1}{\sqrt{1 - \sum_{i=1}^{n-1} x_i^2}} . \tag{C.1}$$

Example 7 (*Area of* $\mathcal{S}_n$)**.** The area a_n is twice the area of the "upper hemisphere". Therefore, by (C.1),

$$\begin{aligned} a_n &= 2 \int\limits_{\mathcal{B}_{n-1}} \frac{dx_1 dx_2 \cdots dx_{n-1}}{\sqrt{1 - \sum_{i=1}^{n-1} x_i^2}} , \qquad \text{(C.2)} \\ &= 2 \int_0^1 \frac{dv_{n-1}(r)}{\sqrt{1-r^2}} , \quad \text{using spherical shells of radius } r \text{ and volume } dv_{n-1}(r) , \\ &= 2\, a_{n-1} \int_0^1 \frac{r^{n-2}}{\sqrt{1-r^2}}\, dr , \quad \text{by (30)} . \\ \therefore \quad \frac{a_n}{a_{n-1}} &= \int_0^1 (1-x)^{-1/2}\, x^{(n-3)/2}\, dx , \quad \text{using } x = r^2 , \\ &= B\left(\frac{n-1}{2}, \frac{1}{2}\right) = \frac{\Gamma(\frac{n-1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{n}{2})} \qquad \text{(C.3)} \end{aligned}$$

and a_n can be computed recursively, beginning with $a_2 = 2\pi$, giving (31).

Example 8 (*Volume of* $\mathcal{B}_n$)**.** The volume of the unit ball $\mathcal{B}_n$ can be computed by (30) and (31),

$$v_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2} + 1\right)} , \quad n = 1, 2, \cdots \tag{C.4}$$

Alternatively, the volume v_n can be computed as the limit of the sum of volumes of cylinders, with base $dx_1 \cdots dx_{n-1}$ and height $2\sqrt{1 - \sum_{k=1}^{n-1} x_k^2}$,

$$v_n = 2 \int\limits_{\mathcal{B}_{n-1}} \sqrt{1 - \sum_{k=1}^{n-1} x_k^2}\ dx_1 \cdots dx_{n-1} \tag{C.5}$$

a routine integration, similar to Example 7.

RUTCOR–Rutgers Center for Operations Research, Rutgers University, 640 Bartholomew Rd, Piscataway, NJ 08854-8003, USA
E-mail address: bisrael@rutcor.rutgers.edu