Lecture Notes in Mathematics 1617

Vadim Yurinsky

Sums and Gaussian Vectors

Springer

Vadim Yurinsky

Sums and Gaussian Vectors

Springer

Author

Vadim Vladimirovich Yurinsky
Sobolev Institute of Mathematics
Siberian Division of the Russian Academy of Sciences
Universitetsky prospekt, 4
Novosibirsk, 630090, Russia

Library of Congress Cataloging-in-Publication Data

Yurinsky, Vadim, 1945-
Sums and Gaussian vectors / Vadim Yurinsky.
p. cm. -- (Lecture notes in mathematics ; 1617)
Includes bibliographical references (p. -) and index.
ISBN 3-540-60311-5 (soft cover)
1. Gaussian sums. 2. Limit theorems (Probability theory)
I. Title. II. Series: Lecture notes in mathematics (Springer
-Verlag) ; 1617.
QA3.L28 no. 1617
[QA246.8.G38]
510 s--dc20
[519.2'6] 95-20482
CIP

Mathematics Subject Classification (1991): 60B12, 60F05, 60F10

ISBN 3-540-60311-5 Springer-Verlag Berlin Heidelberg New York

Typesetting: Camera-ready $\TeX$ output by the author
SPIN: 10479502 46/3142-543210 - Printed on acid-free paper

Preface

In a traditional course of probability theory, the main result is the central limit theorem — the assertion that Gaussian distributions approximate those of sums of independent random variables. Since its discovery, probabilists have had to work hard, both learning calculus and inventing their own tools, before the view of the effects that appear in summing real random variables became reasonably clear.

The more recent shift to study of distributions in "very high" or infinite dimensions brought new discoveries and disappointments.

A fact that can safely be included into either category is that many results for sums of infinite-dimensional random vectors are quite similar to their real-line counterparts, and can be obtained by extremely simple means.

This book is an exposition of this homely part of probability in infinite dimensions — inequalities and asymptotic expressions for large deviation probabilities, normal and related approximations for distributions of sums — all of them considered as preparations for the feat of getting a very general CLT with an estimate of convergence rate.

Naturally, something is lost upon the passage from one to infinitely many dimensions — distributions in high dimensions can develop all real-line pathologies as well as some unthinkable in the classical context. Thus, most results concerning convergence rates in the CLT become fairly bulky if attention is paid to details. This makes the choice between transparency and completeness of exposition even less easy.

The compromise attempted here is to provide a reasonably detailed view of the ideas that have already gained a firm hold, to make the treatment as unified as possible, and sacrifice some of the details that do not fit into the scheme or tend to inflate the text beyond reasonable limits. The price is that such a selection inevitably results biased, and one of the sacrifices was the refined CLT itself. Bibliographic commentary is intended as a partial compensation of bias.

Most of the text of this book was written in Novosibirsk at the *Institute of Mathematics* of Siberian Division of the Russian Academy of Sciences. At the final stage, the author stayed for some time at Departamento de Matemáticas de *Universidad de Oviedo* as a scholarship holder of the *FICYT* (Fundación por el Fomento en Asturias de Investigación Científica Aplicada Y Tecnología). I am sincerely grateful to these two institutions for the opportunity to work without haste they gave me.

I am sincerely grateful to my colleagues, in Novosibirsk, Moscow, Vilnius, Kiev, and many other places, who shared their ideas and created the intellectual stimuli indispensable for any mathematical research, and I profit of the opportunity to thank specially the librarians of the Institute of Mathematics who manage to keep our Novosibirsk library fairly complete despite all the difficulties.

I express my gratitude to Professors Yu.V.Prokhorov and V.V.Sazonov for advice and encouragement.

Contents

Chapter 1

Gaussian Measures in Euclidean Space

This chapter exposes some theorems about the distributions of convex functions with a Gaussian random vector for argument. The results included are mainly inequalities for specific characteristics of such laws. Some of them will later serve to derive similar estimates in infinite dimensions.

The terms "normal" and "Gaussian" (distribution, random variable, etc.) are used as synonyms. The former usually refers to the real line, and the latter to the case of several dimensions.

1.1 Preliminaries

This section is a reminder of basic facts concerning Gaussian distributions in Euclidean spaces.

1.1.1 Standard Normal Distribution

On the real line, the density and distribution function (DF) of the standard normal distribution are

$$\varphi(x) = \exp\left\{-\tfrac{1}{2}x^2\right\}/\sqrt{2\pi}, \quad \Phi(x) = \int_{-\infty}^{x} \varphi(y)dy. \tag{1.1.1}$$

Its characteristic function (CF) is $\int_{-\infty}^{\infty} e^{itx}\varphi(x)dx = \exp\{-t^2/2\}$.

For $x \to \infty$, the behavior of the standard normal DF is described by the asymptotic relation

$$1 - \Phi(x) = [1 + \mathcal{O}(1/x^2)]\varphi(x)/x. \tag{1.1.2}$$

There are similar expressions for its derivatives:

$$\Phi'(x) = x\,[1 - \Phi(x)]\,(1 + o(1))\,, \quad \Phi''(x) = -x^2\,[1 - \Phi(x)]\,(1 + o(1))\,. \tag{1.1.3}$$

The following estimates for integrals with respect to the standard normal distribution (1.1.1) are used in some calculations below.

Lemma 1.1.1 *For* $t \in \mathbf{R}$ *and* $l = 1, 2$,

$$I_l(t) \equiv \int_t^\infty (x-t)^l d\Phi(x) \le (1+t_-)^l .$$

Proof of Lemma 1.1.1. The inequality $I_2(t) \le I_2(0) = \frac{1}{2}$ is evident for $t \ge 0$ since the derivative $I_2'(t)$ is nonpositiive.

If $t < 0$, then

$$I_2(t) \le \int_{-\infty}^\infty \left[x^2 + 2|x|t_- + t_-^2\right] d\Phi(x) \le (1+t_-)^2 .$$

The estimate for $I_1(t)$ is derived similarly. ○

Lemma 1.1.2 *If* $\tau > t \ge 0$, *then*

$$\int_\tau^\infty (x-t)^l d\Phi(x) \le l!(\tau-t)^l [1-\Phi(\tau)] \sum_{m=0}^{l} \frac{\alpha^m}{(l-m)!}, \quad \alpha = \frac{1}{\tau(\tau-t)}.$$

Proof of Lemma 1.1.2. If $\tau > 0$, then the functions

$$p_1(x) = \tau \exp\{-\tau(x-\tau)\}, \quad p_2(x) = \varphi(x)/[1-\Phi(\tau)]$$

are both densities of distributions concentrated on the half-line $[\tau, \infty)$.

The ratio $p_1(x)/p_2(x)$ is increasing and the integrals of both functions over $[\tau, \infty)$ equal 1, so there is a number $A \ge \tau$ such that $p_1(x) \ge p_2(x)$ for $x > A$ and $p_1(x) \le p_2(x)$ for $x < A$. Consequently, for each increasing function f

$$\int_\tau^\infty f(x)[p_1(x) - p_2(x)]\,dx = \int_\tau^\infty [f(x) - f(A)][p_1(x) - p_2(x)]\,dx \ge 0,$$

and there is inequality

$$\begin{aligned} \int_\tau^\infty f(x) d\Phi(x) &\le [1-\Phi(\tau)] \int_\tau^\infty f(x)\tau e^{-\tau(x-\tau)}dx \\ &= [1-\Phi(\tau)] \int_0^\infty f(\tau + x/\tau)\, e^{-x}\,dx. \end{aligned}$$

It is easy to calculate the integral if $f(x) = (x-t)^l$. The calculation yields the inequality of the lemma. ○

1.1.2 Gaussian Distributions: Basic Definitions

The *standard Gaussian distribution* in $\mathbf{R}^k$, $N_{k,0,I}$, is one with the density

$$n_{k,0,I}(x) = (2\pi)^{-k/2} \exp\left\{-\frac{1}{2}|x|^2\right\}. \tag{1.1.4}$$

A random vector (RV) with distribution (1.1.4) is also called standard Gaussian. Its coordinates are independent random variables (rv's) with distribution (1.1.1).

If the distribution of the RV $\eta \in \mathbf{R}^l$ is $N_{l,0,I}$, then, whatever the choice of a constant vector $a \in \mathbf{R}^k$ and $k \times l$ matrix A, the RV

$$\xi = a + A\eta \in \mathbf{R}^k \tag{1.1.5}$$

has the CF

$$g_{k,a,V}(t) \equiv \mathbf{E}\exp\{i(t,\xi)\} = \exp\left\{i(a,t) - \frac{1}{2}(Vt,t)\right\}, \tag{1.1.6}$$

where $t \in \mathbf{R}^k$ and $V = AA^T$. The covariance matrix of ξ and its expectation are

$$\left\|\mathbf{cov}\left(\xi^{(i)},\xi^{(j)}\right)\right\| = V, \quad \mathbf{E}\xi = a, \tag{1.1.7}$$

where $\mathbf{cov}\left(\xi^{(i)},\xi^{(j)}\right) = \mathbf{E}\xi^{(i)}\xi^{(j)} - \mathbf{E}\xi^{(i)}\mathbf{E}\xi^{(j)}$. Each nonnegative matrix admits some factorization $V = AA^T$ with $l \geq k$. Consequently, the right-hand side of (1.1.6) defines the CF of a probability law in $\mathbf{R}^k$ for any a and $V \geq 0$.

A ***Gaussian distribution*** in $\mathbf{R}^k$ is a distribution with CF (1.1.6). It is denoted by $N_{k,a,V}$. Formula (1.1.5) shows that a RV with an arbitrary Gaussian distribution can be obtained from a standard Gaussian one with same or greater number of coordinates by an affine transform. This proves, in particular, that for each Gaussian RV

$$\mathbf{E}\exp\left\{h\left|\xi\right|^2\right\} < \infty \text{ if } |h| < h_0(V).$$

The parameters a and V in (1.1.6) are the expectation and covariance matrix of the distribution. The Gaussian distribution $N_{k,a,V}$ is *nondegenerate* if $V > 0$.

Theorem 1.1.1 *A nondegenerate Gaussian distribution $N_{k,a,V}$ has density*

$$n_{k,a,V}(x) = \frac{1}{\sqrt{\det(2\pi V)}}\exp\left\{-\frac{1}{2}\left(V^{-1}[x-a],[x-a]\right)\right\}. \tag{1.1.8}$$

If the matrix V is degenerate, the Gaussian distribution $N_{k,a,V}$ is concentrated in the affine manifold $L = a + V^{1/2}\mathbf{R}^k$. It is absolutely continuous with respect to the Lebesgue measure in this Euclidean space (whose metric is inherited from the original space). This is easily verified, e.g., using (1.1.5).

Proof of Theorem 1.1.1. The right-hand side of (1.1.6) is summable if $V > 0$, so the distribution corresponding to this CF does indeed have the continuous density

$$n_{k,a,V}(x) = (2\pi)^{-k}\int \exp\left\{-i(x-a,t) - \tfrac{1}{2}(Vt,t)\right\}\,dt.$$

By changing variables to $y = V^{1/2}t$ and $z = V^{-1/2}(x-a)$, the integral is transformed into

$$n_{k,a,V}(x) = n_{k,0,I}(z)/\sqrt{\det(V)} = g_{k,0,I}(z)/\sqrt{\det(2\pi V)}.$$

This proves the theorem. ○

1.1.3 Characterization Theorems

Gaussian distributions are infinitely divisible and stable. It follows from the definition of a multidimensional Gaussian distribution that a RV is Gaussian if and only if its projection onto each one-dimensional subspace is Gaussian. This circumstance allows one to extend known theorems concerning characterization of Gaussian distributions on the real line to multidimensional Euclidean spaces.

Theorem 1.1.2 *Let $\xi,\ \eta \in \mathbf{R}^k$ be independent RV's.*
If $\xi + \eta$ is Gaussian, then both ξ and η are Gaussian as well.

Theorem 1.1.3 *Let ξ and η be i.i.d. RV's. Put*

$$\xi(a) = \cos(a)\xi + \sin(a)\eta,\ \eta(a) = -\sin(a)\xi + \cos(a)\eta.$$

a) *If, for some $a \in (0, \pi/2)$, the RV's $\xi(a)$ and $\eta(a)$ are independent and have identical distributions, then ξ and η are centered Gaussian RV's: $\mathcal{L}(\xi) = \mathcal{L}(\eta) = N_{k,0,V}$ with some matrix $V \geq 0$.*

b) *If ξ and η are independent and have distribution $N_{k,0,V}$, then $\xi(a)$ and $\eta(a)$ are, for each value of a, also independent with the same distribution: $\mathcal{L}(\xi) = \mathcal{L}(\eta) = N_{k,0,V}$.*

1.1.4 Monotonicity in Covariances

The next theorem is a special case of the so-called Slepian inequality. It shows that for a Gaussian RV the dependence of the *joint DF function of coordinates*

$$\mathbf{P}\{\xi < x\} = \mathbf{P}\left(\bigcap_{j=1}^{k} \{\xi^{(j)} < x^{(j)}\}\right)$$

on its covariance matrix is, in a sense, monotone.

The CF inversion formula yields the following expression for the DF of a Gaussian RV: if $\mathcal{L}(\xi) = N_{k,a,V}$, then

$$N_{k,a,V}(x) = \mathbf{P}\{\xi < x\} = \int_{y<x} \left[(2\pi)^{-k} \int e^{i(t,y)} g_{k,a,V}(t)dt\right] dy. \qquad (1.1.9)$$

It follows from this equality that, for x fixed, the DF is a smooth function in its parameters a and V on the set $\left\{\langle a, V\rangle \in \mathbf{R}^k \times \left(\mathbf{R}^k \otimes \mathbf{R}^k\right) : V > 0\right\}$.

Lemma 1.1.3 *Consider the Gaussian distribution $N_{k,a,\widetilde{V}}$ with the covariance matrix*

$$\widetilde{V} = \widetilde{V}(u) = V + uC,\ u \in \mathbf{R}.$$

where $V > 0$ and the real square matrix $C = (C_{ij})$ is symmetrical with zero diagonal elements: $c_{ii} = 0,\ i = \overline{1,k}$. The DF of this distribution satisfies the relations

$$\left.\frac{d}{du} N_{k,a,\widetilde{V}}(x)\right|_{u=0} = \frac{1}{2} \sum_{l \neq m} c_{lm} \int_{A_{lm}(x)} n_{k,a,V}(y)\, \mathrm{mes}_{lm}(dy) \geq 0,$$

where mes_{lm} *is the Lebesgue measure in the* $(k-2)$*-dimensional orthant*

$$A_{lm}(x) = \left\{y:\ y^{(j)} < x^{(j)},\ j \neq l, m;\ y^{(l)} = x^{(l)},\ y^{(m)} = x^{(m)}\right\}.$$

Proof of Lemma 1.1.3. CF (1.1.6) decays rapidly as $|t| \to \infty$. Hence for small values of u it is possible to interchange the order of differentiation and integration. It follows that in a neighborhood of zero

$$\frac{d}{du} N_{k,a,\widetilde{V}}(x) = \frac{-1}{2(2\pi)^k} \int_{y<x} \left(\sum_{l \neq m} c_{lm} \int t^{(l)} t^{(m)}\, g_{k,a,\widetilde{V}}(t)\, e^{-i(t,y)}\, dt \right) dy.$$

The inner integral in the right-hand side can be represented as a derivative with respect to $y^{(l)}$ and $y^{(m)}$: indeed,

$$\begin{aligned} \int t^{(l)} t^{(m)}\, g_{k,a,\widetilde{V}}(t)\, e^{-i(t,y)}\, dt &= -\frac{\partial^2}{\partial y^{(l)} \partial y^{(m)}} \int g_{k,a,\widetilde{V}}(t)\, e^{-i(t,y)}\, dt \\ &= -(2\pi)^k \frac{\partial^2 n_{k,a,\widetilde{V}}}{\partial y^{(l)} \partial y^{(m)}}(y). \end{aligned}$$

To obtain the assertion of the lemma, it suffices to integrate the last identity in variables $y^{(l)}$, $y^{(m)}$ and put $u = 0$. ○

Theorem 1.1.4 *Assume that* $\mathcal{L}(\xi) = N_{k,a,V}$ *and* $\mathcal{L}(\eta) = N_{k,a,W}$. *If the covariance matrices of these Gaussian distributions satisfy the relations*

$$V_{jj} = W_{jj}, \quad V_{ij} \leq W_{ij},\ i \neq j,$$

then for all $x \in \mathbf{R}^k$

$$\mathbf{P}\{\xi < x\} \leq \mathbf{P}\{\eta < x\}.$$

Proof of Theorem 1.1.4. If both covariance matrices are nondegenerate, the inequality of the theorem follows from Lemma 1.1.3. Indeed, in this case the matrix $\widetilde{V} = uW + (1-u)V$ is also nondegenerate for all $u \in [0,1]$, and $C = W - V$ satisfies the conditions of this lemma, so $(d/du) N_{k,a,\widetilde{V}(u)}(x) \geq 0$.

If at least one of the covariance matrices degenerates, one can first derive the inequality of the theorem for the distributions with covariance matrices $V + \varepsilon I$ and $W + \varepsilon I$ assuming that $\varepsilon > 0$.

As $\varepsilon \to 0$, the distributions corresponding to these latter matrices weakly converge to the original ones. This convergence yields the inequality of the theorem at all continuity points of the DF's considered. The set of these points is dense in $\mathbf{R}^k$. Thus, one more passage to the limit shows that the inequality holds everywhere. ○

1.1.5 Conditional Distributions and Projections

For Gaussian rv's, independence and absence of correlations are equivalent.

Theorem 1.1.5 *Let the joint distribution of the real rv's ξ and η be a Gaussian one. Then ξ and η are independent if and only if they are uncorrelated, i.e., if* $\mathbf{cov}(\xi, \eta) = 0$.

Proof of Theorem 1.1.5 reduces to calculating the joint CF:

$$\mathbf{E}\exp\{is\xi + it\eta\} = \exp\left\{is\mathbf{E}\xi + it\mathbf{E}\eta - \tfrac{1}{2}s^2\mathbf{D}\xi - \tfrac{1}{2}t^2\mathbf{D}\eta - st\,\mathbf{cov}(\xi,\eta)\right\}.$$

The right-hand side of this equality splits into the product of CF's of ξ and η only if the covariance equals zero. ○

One more result of this kind is

Theorem 1.1.6 *Let $\xi \in \mathbf{R}^k$ be a Gaussian RV with the distribution $N_{k,a,V}$ and $b \in \mathbf{R}^k$ a constant vector. Assume, moreover, that $(Vb, b) > 0$, and put $\eta = (\xi, b)$.*

The conditional distribution of ξ given $\eta = y$ is Gaussian with expectation and covariance matrix

$$m(y) = a - (Vb, b)^{-1}[y - (a, b)]\,Vb, \quad \Sigma = V - (Vb, b)^{-1}Vb(Vb)^T,$$

i.e., $\mathbf{P}\{\xi \in A\,|\eta = y\} = N_{k,m(y),\Sigma}(A)$ *for each measurable subset* $A \subset \mathbf{R}^k$.

Proof of Theorem 1.1.6. The RV ξ can be represented in the form $\xi = \eta B + X$, where the coordinates of X are uncorrelated with the real rv η. To obtain a decomposition of this kind it suffices to set $B = (Vb, b)^{-1}\,Vb$ and $X = \xi - \eta B$.

Apply (1.1.6) to compute $\mathbf{E}\exp\{i(\lambda, X) + i(\mu, \eta B)\}$, the CF of the $2k$-dimensional joint distribution. It splits into the product of marginal CF's, those of X and ηB. Hence these RV's are independent. The same calculation shows they are also Gaussian, and a little more work yields the covariance matrix and expectation: thus, given $\eta = y$ the conditional expectation of ξ equals $\mathbf{E}X + yB$, etc. ○

The following special case will be used later on. Let $\xi = \left(\xi^{(j)},\ j = \overline{1,k}\right)$ be a Gaussian RV. For the conditional distribution of $\left(\xi^{(j)},\ j = \overline{2,k}\right) \in \mathbf{R}^{k-1}$ given $\xi^{(1)} = y$, the coordinates of the expectation are, by Theorem 1.1.6,

$$m^{(j)} = \mathbf{E}\xi^{(j)} + \frac{\mathbf{cov}\left(\xi^{(1)}, \xi^{(j)}\right)}{\mathbf{cov}\left(\xi^{(1)}, \xi^{(1)}\right)}\left(y - \mathbf{E}\xi^{(1)}\right), \tag{1.1.10}$$

and the covariance matrix has the elements

$$v'_{ij} = \mathbf{cov}\left(\xi^{(i)}, \xi^{(j)}\right) - \frac{\mathbf{cov}\left(\xi^{(1)}, \xi^{(i)}\right)\mathbf{cov}\left(\xi^{(1)}, \xi^{(j)}\right)}{\mathbf{cov}\left(\xi^{(1)}, \xi^{(1)}\right)}. \tag{1.1.11}$$

1.1.6 Laplace Transform for DF of Squared Norm

Let $\xi \in \mathbf{R}^k$ be a Gaussian RV with the distribution $N_{k,a,V}$ whose covariance matrix V has the eeigenvalues

$$\sigma_1^2 = \cdots = \sigma_\nu^2 > \sigma_{\nu+1}^2 \geq \cdots \geq \sigma_k^2,$$

where ν is the multiplicity of the principal eigenvalue.

Theorem 1.1.7 *For each complex number z such that $2\sigma_1^2 \,\mathbf{Re}\, z < 1$*

$$\mathbf{E}\exp\left\{z\,|\xi|^2\right\} = \exp\left\{z\left(\left[I - 2zV\right]^{-1} a, a\right)\right\} \prod_{j=1}^{k} 1/\sqrt{1 - 2z\sigma_j^2}.$$

The branch of the complex-valued square root in the right-hand side is selected so as to satisfy the condition $\sqrt{1 - 2z\sigma_j^2} > 0$ for $\mathbf{Im}\, z = 0$.

Proof of Theorem 1.1.7. Let Q be an orthogonal matrix which transforms V to its diagonal form: $Q^T V Q = \left(\sigma_j^2 \delta_{ij}\right)$. According to (1.1.5)–(1.1.6), the distribution $N_{k,a,V}$ is that of the RV $a + QD\eta$, where $D = (\sigma_j \delta_{ij})$ is a diagonal matrix and η a standard Gaussian RV. Hence

$$\mathbf{E}\exp\left\{z\,|\xi|^2\right\} = \mathbf{E}\exp\left\{z\,|\bar{a} + D\eta|^2\right\} = \prod_{j=1}^{k} I(\bar{a}^{(j)}, \sigma_j, z),$$

where $\bar{a} = Q^T a = \left(\bar{a}^{(j)}\right) \in \mathbf{R}^k$, $|\bar{a}| = |a|$, and

$$I(\bar{a}, \sigma, z) = \frac{1}{\sqrt{2\pi}} \exp\left\{z\left|\bar{a}^2\right|\right\} \int_{-\infty}^{\infty} \exp\left\{-\tfrac{1}{2}\left(1 - 2z\sigma^2\right) u^2 + 2z\sigma\bar{a}\right\} du.$$

The integrand in I is analytical, so the choice of integration path in the complex plane does not influence the value of the integral. If $\mathbf{Re}\left(1 - 2z\sigma^2\right) > 0$, one arrives at the equality

$$\begin{aligned} I(\bar{a}, \sigma, z) &= \frac{1}{\sqrt{2\pi}} \exp\left\{\frac{z\bar{a}^2}{(1 - 2z\sigma^2)}\right\} \int_{\mathrm{Im}\, u=0} \exp\left\{-\frac{1 - 2z\sigma^2}{2}(u - c)^2\right\} du \\ &= \exp\left\{\frac{z\bar{a}^2}{1 - 2z\sigma^2}\right\} / \sqrt{1 - 2z\sigma_j^2}. \end{aligned}$$

The equality of the theorem now results from the relation

$$\sum_{j=1}^{k} \frac{\left(\bar{a}^{(j)}\right)^2}{1 - 2z\sigma_j^2} = \left(\left[I - 2zD^2\right]^{-1} \bar{a}, \bar{a}\right) = \left(\left[I - 2zV\right]^{-1} a, a\right). \bigcirc$$

1.2 Extremal Properties of Half-Spaces

1.2.1 Isoperimetric Property of Sphere

The ***uniform distribution*** on the sphere $S_r \subset \mathbf{R}^k$ is defined by the equality

$$\sigma_r(A) = |S_r|^{-1} \int_{A \bigcap \{|x|=r\}} ds(x), \tag{1.2.1}$$

where $|S_r|$ is the surface area of the sphere (see (A.1.7) and (A.1.8)).

Distribution (1.2.1) can be considered both as a measure on Borel subsets of $\mathbf{R}^k$ and as one on subsets of the surface of the sphere. Whenever this does not lead to misunderstanding, no special effort is spent on discerning between the two:

$$\sigma_r(A) = \sigma_r\left(A \bigcap S_r\right), \quad A \in \mathcal{B}\left(\mathbf{R}^k\right); \quad \sigma_r\left(\mathbf{R}^k\right) = \sigma_r(S_r) = 1. \tag{1.2.2}$$

For a half-space

$$H_{e,\rho} = \{x : (x,e) \le \rho\}, \; H^c_{e,\rho}\{x : (x,e) > \rho\}, \tag{1.2.3}$$

the value of $\sigma_r(H_{e,\rho})$ does not depend on the direction of the unit vector e, $|e| = 1$. For instance, $\sigma_r(H_{e,0}) = \sigma_r\left(H^c_{e,0}\right) = \frac{1}{2}$, whatever the choice of e. If $\rho > 0$, then $\sigma_r(H_{e,\rho}) > \frac{1}{2}$ and $\sigma_r\left(H^c_{e,\rho}\right) < \frac{1}{2}$.

Half-spaces possess an important extremal property with respect to the uniform law on sphere. It is expressed by the so-called *isoperimetric inequality* for sphere which is described below in Proposition 1.2.1. Some new notation is necessary to state it.

On the sphere S_r, topology can be defined using the *geodesic distance* $\rho(x,y)$ equal to the length of the shortest geodesic connecting x and y. A geodesic neighborhood of a set $A \subset S_r$ is (cf. (A.1.1))

$$\mathcal{U}_r[A,\delta] = \{x \in S_r : \rho(x,A) < \delta\}, \quad \rho(x,A) = \inf_{y \in A} \rho(x,y). \tag{1.2.4}$$

Proposition 1.2.1 *Let A be an arbitrary measurable subset of $\mathbf{R}^k$. If*

$$\sigma_r(A) \ge \sigma_r(H)$$

for some subspace $H = H_{e,\rho}$, $e \in \mathbf{R}^k$, $\rho > 0$, then the values of the uniform distribution on geodesic neighborhoods of these sets are related through the inequality

$$\forall\, \delta > 0 \;\; \sigma_r\left(\mathcal{U}_r\left[A \bigcap S_r, \delta\right]\right) \ge \sigma_r\left(\mathcal{U}_r\left[H \bigcap S_r, \delta\right]\right). \tag{1.2.5}$$

This proposition will not be proved here (see Appendix B for references).

1.2.2 Constructing Gaussian Law from Uniform Ones

The standard Gaussian distribution can be approximated by marginal distributions of uniform laws (1.2.1) on spheres in very high dimensions.

Below, the space $\mathbf{R}^k$ is considered as a k-dimensional subspace of the Euclidean space $\mathbf{R}^D$ with large (and later infinitely growing) number of dimensions. The corresponding orthogonal projection is denoted by $\mathcal{P} : \mathbf{R}^D \to \mathbf{R}^k$, and $\sigma_{r,D}$ is the uniform distribution (1.2.1) on the sphere of the larger space, $S^{(D)}_r = \left\{x \in \mathbf{R}^D : |x| = r\right\}$.

Lemma 1.2.1 *The distribution in $\mathbf{R}^k$ defined by the formula*

$$\nu_{r,D}(A) \equiv \sigma_{r,D}\left(\left\{z \in S^{(D)}_r : \mathcal{P}z \in A\right\}\right)$$