Записки научных
семинаров ПОМИ
Том 457, 2017 г.

S. G. Bobkov, G. P. Chistyakov, F. Götze

GAUSSIAN MIXTURES AND NORMAL APPROXIMATION FOR V. N. SUDAKOV'S TYPICAL DISTRIBUTIONS

Abstract. We derive a general upper bound on the distance of the standard normal law to typical distributions in V. N. Sudakov's theorem (in terms of the weighted total variation).

Dedicated to the memory of Vladimir Nikolayevich Sudakov

§1. Introduction

Let $X = (X_1, \dots, X_n)$ be a random vector in $\mathbf{R}^n$ with finite second moment, and let

$$S^{n-1} = \{\theta = (\theta_1, \dots, \theta_n) \in \mathbf{R}^n : \theta_1^2 + \dots + \theta_n^2 = 1\}$$

denote the unit sphere which we equip with the uniform probability measure σ_{n-1}.

In general, the distribution functions $F_\theta(x) = \mathbf{P}\{S_\theta \leqslant x\}$ $(x \in \mathbf{R})$ of linear forms

$$S_\theta = \theta_1 X_1 + \dots + \theta_n X_n, \qquad \theta \in S^{n-1},$$

essentially depend on the parameter θ. Nevertheless, according to the celebrated result by Sudakov of 1978 [20], if n is large, and if the covariance matrix of X has a bounded spectral radius, then F_θ's concentrate around a certain typical distribution function F (for most of θ in the sense of σ_{n-1}). The latter function may actually be defined explicitly as the mean

$$F(x) = \int_{S^{n-1}} F_\theta(x)\, d\sigma_{n-1}(\theta). \tag{1.1}$$

Key words and phrases: Sudakov's typical distributions, normal approximation.
Bobkov: Research was partially supported by NSF grant DMS-1612961.
Chistyakov, Götze: Research was partially supported by SFB 1283 at Bielefeld University.

This remarkable observation, to which Sudakov returned several times later on (cf. e.g. [19,21]), has become a starting point for subsequent investigations by many researchers. And indeed, his theorem has a rather universal range of applicability in contrast with the classical scheme of summation of independent random variables. The problem of concentration of F_θ has various interesting aspects, and we do not discuss it here. Let us only mention the papers by Nagaev [17] and von Weizsäcker [22] who considered summation and averaging with coefficients over the rescaled Gaussian measure (instead of σ_{n-1}). The paper [3] dealt with coefficients of the form $\theta_k = \pm 1/\sqrt{n}$ and averaging with respect to the rescaled Bernoulli measure; some other related models were studied in [4,6,7]. For the problem of rates of approximation, and results in the case where the distribution of X has convexity properties, see also [1,2,5,8,9,11–14,16,18].

It was already emphasized in [20] that the typical distribution F in (1.1) may be approximated by a mixture of centered Gaussian measures on the line. Indeed, the rotational invariance of the measure σ_{n-1} implies that

$$F(x) = \mathbf{P}\{\rho Z_n \leqslant x\},$$

where

$$\rho^2 = \frac{X_1^2 + \dots + X_n^2}{n} \qquad (\rho \geqslant 0),$$

and where the random variable Z_n is independent of ρ and has the same distribution as $\sqrt{n}\theta_1$ under σ_{n-1}. Since Z_n is nearly standard normal, F is therefore close to the distribution of ρZ with $Z \sim N(0,1)$ independent of ρ.

In particular, F itself is approximately normal, if and only if ρ is almost a constant, which means a kind of the law of large numbers for the sequence X_k^2. This property – that the distribution of ρ is concentrated around a point (in a weak sense) is of course true in case of independent components X_k's (under a mild moment assumption), but it continues to hold in many other situations allowing dependence between X_k. To quantify the assertion about the closeness of F to the standard normal distribution function

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2}\, dy, \qquad x \in \mathbf{R},$$

we derive a simple general bound in terms of the variance of ρ. Note that the second moment of F is equal to $\mathbf{E}\rho^2$, so a normalization condition on this moment is desirable.

Theorem 1.1. *Suppose that* $\mathbf{E}\rho^2 = 1$*. With some absolute constant* $c > 0$*, we have for all* $n \geqslant 1$*,*

$$\int_{-\infty}^{\infty} (1 + x^2)\, |F - \Phi|(dx) \,\leqslant\, c\,\Big(\frac{1}{n} + \mathrm{Var}(\rho)\Big). \tag{1.2}$$

Here the positive measure $|F - \Phi|$ denotes the variation in the sense of measure theory, and the left integral represents the weighted total variation of $F - \Phi$. In particular, we have a similar bound on the usual total variation distance between F and Φ.

In applications, it might be more convenient to use an elementary bound $\mathrm{Var}(\rho) \leqslant \mathrm{Var}(\rho^2)$, cf. (3.2) below, in order to further estimate the right-hand side of (1.2). For example, if the random variables X_k are pairwise independent and have bounded 4-th moments $\mathbf{E}X_k^4$, then $\mathrm{Var}(\rho^2)$ is of order $1/n$, so that (1.2) yields a $\frac{1}{n}$-rate of normal approximation for the total variation and thus for the Kolmogorov distance $\Delta = \sup_x |F(x) - \Phi(x)|$ as well. Another wide class of probability distributions with this property (for X) is described by those that satisfy a Poincaré-type inequality

$$\lambda_1 \mathrm{Var}(u(X)) \leqslant \mathbf{E}\, |\nabla u(X)|^2,$$

where a positive constant $\lambda_1 = \lambda_1(X)$ serves for all bounded smooth functions u on $\mathbf{R}^n$. The appearance of the weight $1 + x^2$ on the left of (1.2) allows one to make a similar conclusion about the L^p-distances between the distribution functions F and Φ.

§2. Mixtures of Centered Gaussian Measures

As a first natural step towards the proof of Theorem 1.1, let us consider the normal approximation for general mixtures of normal distributions. Denote by Φ_ρ the distribution function of the random variable $Z(\rho) = \rho Z$, where $Z \sim N(0,1)$ is independent of a random variable $\rho \geqslant 0$. That is,

$$\Phi_\rho(x) = \mathbf{P}\{\rho Z \leqslant x\} = \mathbf{E}\,\Phi(x/\rho), \qquad x \in \mathbf{R}.$$

We start with estimation of the weighted total variation distance between the distributions Φ_ρ and Φ.

Proposition 2.1. *If* $\mathbf{E}\,\rho^2 = 1$*, then*

$$\int_{-\infty}^{\infty} (1 + x^2)\, |\Phi_\rho - \Phi|(dx) \,\leqslant\, c\,\mathrm{Var}(\rho), \tag{2.1}$$

where c is an absolute constant.

Proof. When $\rho = t$ is a positive constant, Φ_t represents the normal distribution function with mean zero and standard deviation $t > 0$, thus with density

$$\varphi_t(x) = \frac{1}{t}\,\varphi(x/t), \quad x \in \mathbf{R},$$

where $\varphi(x) = \frac{1}{\sqrt{2\pi}}\,e^{-x^2/2}$ is the standard normal density.

For a fixed number x, let us expand the function $u(t) = \varphi_t(x)$ according to the Taylor formula up to the quadratic term at the point $t_0 = 1$. We have $u(t_0) = \varphi(x)$,

$$u'(t) = (t^{-4}x^2 - t^{-2})\,\varphi(x/t), \qquad u'(t_0) = (x^2 - 1)\,\varphi(x),$$

and

$$u''(t) = (t^{-7}x^4 - 5t^{-5}x^2 + 2t^{-3})\,\varphi(x/t) = t^{-3}\,\psi(x/t),$$

where $\psi(z) = (z^4 - 5z^2 + 2)\,\varphi(z)$. Therefore, using the integral Taylor formula

$$u(t) = u(t_0) + u'(t_0)(t - t_0) + (t - t_0)^2 \int_0^1 u''((1-s) + st)\,(1-s)\,ds, \tag{2.2}$$

we get

$$\varphi_t(x) - \varphi(x) = (t-1)\,(x^2-1)\,\varphi(x) + (t-1)^2 \int_0^1 \big((1-s)+st\big)^{-3}\,\psi\Big(\frac{x}{(1-s)+st}\Big)\,(1-s)\,ds.$$

We apply this representation with $t = \xi(\omega)$, where ξ is a positive random variable (on some probability space). Thus, Φ_ξ has density $\varphi_\xi = \mathbf{E}\,\varphi_{\xi(\omega)}$ representable as

$$\varphi_\xi(x) - \varphi(x) = (x^2-1)\,\varphi(x)\,\mathbf{E}\,(\xi-1) + \mathbf{E}\,(\xi-1)^2 \int_0^1 \big((1-s)+s\xi\big)^{-3}\,\psi\Big(\frac{x}{(1-s)+s\xi}\Big)\,(1-s)\,ds.$$

Putting

$$R_\xi(x) = \varphi_\xi(x) - \varphi(x) - (x^2-1)\,\varphi(x)\,\mathbf{E}\,(\xi-1),$$

we then get, by Fubini's theorem,

$$\begin{aligned}
&\int_{-\infty}^{\infty} |R_\xi(x)|\,dx \\
&\quad \leqslant \mathbf{E} \int_{-\infty}^{\infty} (\xi-1)^2 \int_0^1 \big((1-s)+s\xi\big)^{-3} \Big|\psi\Big(\frac{x}{(1-s)+s\xi}\Big)\Big|\,(1-s)\,ds\,dx \\
&\quad = c_0\,\mathbf{E}\,(\xi-1)^2 \int_0^1 \big((1-s)+s\xi\big)^{-2}\,(1-s)\,ds
\end{aligned}$$

with $c_0 = \int_{-\infty}^{\infty} |\psi(x)|\,dx$. In particular, if $\xi \geqslant \frac{1}{2}$, the latter integral does not exceed $\int_0^1 \frac{1-s}{(1-\frac{s}{2})^2}\,ds = 4\log 2 - 2 < 1$, so that

$$\int_{-\infty}^{\infty} |R_\xi(x)|\,dx \leqslant c_0\,\mathbf{E}\,(\xi-1)^2,$$

which implies that with some $|\theta| \leqslant 1$

$$\int_{-\infty}^{\infty} |\varphi_\xi(x) - \varphi(x)|\,dx = |\mathbf{E}\xi - 1| \int_{-\infty}^{\infty} |x^2-1|\,\varphi(x)\,dx + \theta c_0\,\mathbf{E}\,(\xi-1)^2.$$

Analogously, the integral $\int_{-\infty}^{\infty} x^2\,|R_\xi(x)|\,dx$ may be bounded from above by

$$\mathbf{E} \int_{-\infty}^{\infty} (\xi-1)^2 \int_0^1 \big((1-s)+s\xi\big)^{-3} x^2 \Big|\psi\Big(\frac{x}{(1-s)+s\xi}\Big)\Big|\,(1-s)\,ds\,dx = c_1\,\mathbf{E}\,(\xi-1)^2,$$

where $c_1 = \frac{1}{2} \int_{-\infty}^{\infty} x^2 \, |\psi(x)| \, dx$. In particular, now without any constraint on ξ,

$$\int_{-\infty}^{\infty} x^2 \, |\varphi_\xi(x) - \varphi(x)| \, dx = |\mathbf{E}\xi - 1| \int_{-\infty}^{\infty} x^2 \, |x^2 - 1| \, \varphi(x) \, dx + \theta c_1 \, \mathbf{E}\,(\xi - 1)^2.$$

The two representations can now be combined to

$$\int_{-\infty}^{\infty} (1 + x^2) \, |\varphi_\xi(x) - \varphi(x)| \, dx \ = \ a \, |\mathbf{E}\xi - 1| + \theta b \, \mathbf{E}\,(\xi - 1)^2,$$

where

$$a = \int_{-\infty}^{\infty} (1 + x^2) \, |x^2 - 1| \, \varphi(x) \, dx, \qquad b = \int_{-\infty}^{\infty} \Big(1 + \frac{1}{2}\, x^2\Big) \, |\psi(x)| \, dx$$

and $|\theta| \leqslant 1$. Let us derive numerical bounds on these absolute constants. If $Z \sim N(0,1)$, then using $(1+x^2)\,|x^2-1| \leqslant 1+x^4$, we have $a \leqslant 1+\mathbf{E}Z^4 = 4$. Since furthermore $|\psi(x)| \leqslant (x^4 + 5x^2 + 2)\varphi(x)$, we also have

$$b \leqslant \mathbf{E}\,(Z^4 + 5Z^2 + 2) + \frac{1}{2}\,\mathbf{E}\,(Z^6 + 5Z^4 + 2Z^2) = 26.$$

Thus,

$$\int_{-\infty}^{\infty} (1 + x^2) \, |\varphi_\xi(x) - \varphi(x)| \, dx \ = \ 4\theta_0 \, |\mathbf{E}\,\xi - 1| + 26\,\theta_1 \, \mathbf{E}\,(\xi - 1)^2 \qquad (2.3)$$

with some $|\theta_i| \leqslant 1$, provided that $\xi \geqslant 1/2$.

Now, consider the general case assuming without loss of generality that $\rho > 0$. Introduce the events $A_0 = \{\rho < 1/2\}$, $A_1 = \{\rho \geqslant 1/2\}$, and put $\alpha_0 = \mathbf{P}(A_0)$, $\alpha_1 = \mathbf{P}(A_1)$, assuming again without loss of generality that $\alpha_0 > 0$. Next we split the distribution Q of ρ into the two components supported on $(0, 1/2)$ and $[1/2, \infty)$ and denote by ρ_0 and ρ_1 some random variables distributed respectively as the normalized restrictions of Q to these regions, so that $\rho_0 < 1/2$ and $\rho_1 \geqslant 1/2$. We thus represent the density of Φ_ρ as the convex mixture of two densities

$$\varphi_\rho = \alpha_0 \varphi_{\rho_0} + \alpha_1 \varphi_{\rho_1}, \qquad (2.4)$$

where

$$\varphi_{\rho_0}(x) = \frac{1}{\alpha_0}\,\mathbf{E}\,\varphi(x/\rho)\,1_{\{\rho\in A_0\}}, \qquad \varphi_{\rho_1}(x) = \frac{1}{\alpha_1}\,\mathbf{E}\,\varphi(x/\rho)\,1_{\{\rho\in A_1\}}.$$

Note that, since $\mathbf{E}\,\rho^2 = 1$, we necessarily have $\mathbf{E}\,\rho \leqslant 1$. On the other hand,

$$\mathrm{Var}(\rho) = 1 - (\mathbf{E}\,\rho)^2 = (1-\mathbf{E}\,\rho)(1+\mathbf{E}\,\rho) \geqslant 1 - \mathbf{E}\,\rho.$$

Hence $|\mathbf{E}\,\rho - 1| \leqslant \mathrm{Var}(\rho)$, and as a consequence,

$$\mathbf{E}\,(\rho-1)^2 = 2\,(1-\mathbf{E}\,\rho) \leqslant 2\,\mathrm{Var}(\rho). \tag{2.5}$$

In particular, since $\rho < 1/2$ implies $(\rho-1)^2 > 1/4$, we have, by Chebyshev's inequality,

$$\alpha_0 = \mathbf{P}\{\rho < 1/2\} \leqslant 8\,\mathrm{Var}(\rho). \tag{2.6}$$

Now, by the previous step (2.3) with $\xi = \rho_1$,

$$\int_{-\infty}^{\infty} (1+x^2)\,|\varphi_{\rho_1}(x) - \varphi(x)|\,dx \;=\; 4\,\theta_0\,|\mathbf{E}\,\rho_1 - 1| + 26\,\theta_1\,\mathbf{E}\,(\rho_1-1)^2. \tag{2.7}$$

On the other hand,

$$\int_{-\infty}^{\infty} (1+x^2)\,\varphi_{\rho_0}(x)\,dx \;=\; 1 + \mathbf{E}\,\rho_0^2 \;\leqslant\; \frac{5}{4},$$

so that

$$\int_{-\infty}^{\infty} (1+x^2)\,|\varphi_{\rho_0}(x) - \varphi(x)|\,dx \;\leqslant\; \frac{13}{4}.$$

From (2.4), (2.6) and (2.7), we now get that

$$\int_{-\infty}^{\infty} (1+x^2)\,|\varphi_{\rho}(x) - \varphi(x)|\,dx \;\leqslant\; 26\,\mathrm{Var}(\rho) + 4\,|\mathbf{E}\,\rho_1 - 1| + 26\,\mathbf{E}\,(\rho_1-1)^2. \tag{2.8}$$

It remains to estimate the last two expectations. First suppose that $\mathrm{Var}(\rho) \leqslant 1/16$, so that, by (2.6), $\alpha_0 \leqslant \frac{1}{2}$ and $\alpha_1 \geqslant \frac{1}{2}$. By definition,

$$\mathbf{E}\,\rho_1 = \frac{1}{\alpha_1}\,\mathbf{E}\,\rho\,1_{\{\rho\in A_1\}} = \frac{1}{\alpha_1}\,\big(\mathbf{E}\,\rho - \mathbf{E}\,\rho\,1_{\{\rho\in A_0\}}\big),$$

hence

$$\mathbf{E}\,\rho_1 - 1 = \frac{1}{\alpha_1}\,\big(\mathbf{E}\,(\rho-1) - \mathbf{E}\,(\rho-1)\,1_{\{\rho\in A_0\}}\big).$$

By Cauchy's inequality and applying (2.5) and (2.6),

$$\begin{aligned}\big|\mathbf{E}\,(\rho-1)\,1_{\{\rho\in A_0\}}\big| &\leqslant \mathbf{E}\,|\rho-1|\,1_{\{\rho\in A_0\}} \\ &\leqslant \big(\mathbf{E}\,(\rho-1)^2\big)^{1/2}\,\alpha_0^{1/2} \leqslant 4\,\mathrm{Var}(\rho).\end{aligned} \tag{2.9}$$

Hence

$$|\mathbf{E}\,\rho_1 - 1| \ \leqslant\ \frac{1}{\alpha_1}\Big(|\mathbf{E}\,\rho-1| + |\mathbf{E}\,(\rho-1)\,1_{\{\rho\in A_0\}}|\Big) \ \leqslant\ 10\,\mathrm{Var}(\rho). \tag{2.10}$$

Similarly,

$$\mathbf{E}\,\rho_1^2 = \frac{1}{\alpha_1}\,\mathbf{E}\,\rho^2\,1_{\{\rho\in A_1\}} = \frac{1}{\alpha_1}\,\big(1-\mathbf{E}\,\rho^2\,1_{\{\rho\in A_0\}}\big),$$

so, using $\rho\leqslant 1/2$ on A_0 and applying (2.9), we have

$$\begin{aligned}\mathbf{E}\,\rho_1^2 - 1 &= \frac{1}{\alpha_1}\,\mathbf{E}\,(1-\rho^2)\,1_{\{\rho\in A_0\}} = \frac{1}{\alpha_1}\,\mathbf{E}\,(1-\rho)\,(1+\rho)\,1_{\{\rho\in A_0\}} \\ &\leqslant \frac{3}{2\alpha_1}\,\mathbf{E}\,|1-\rho|\,1_{\{\rho\in A_0\}} \leqslant \frac{3}{2\alpha_1}\cdot 4\,\mathrm{Var}(\rho) \ \leqslant\ 12\,\mathrm{Var}(\rho).\end{aligned}$$

Writing $\mathbf{E}\,(\rho_1-1)^2 = (\mathbf{E}\,\rho_1^2-1) - 2\,\mathbf{E}\,(\rho_1-1)$ and applying (2.10), these estimates yield

$$\mathbf{E}\,(\rho_1-1)^2 \ \leqslant\ 12\,\mathrm{Var}(\rho) + 20\,\mathrm{Var}(\rho) \ =\ 32\,\mathrm{Var}(\rho).$$

It remains to use this bound together with (2.10) in (2.8) in order to arrive at the desired estimate (2.1), i.e.,

$$\int_{-\infty}^{\infty}(1+x^2)\,|\varphi_\rho(x)-\varphi(x)|\,dx \ \leqslant\ c\,\mathrm{Var}\,(\rho),$$

with the constant $c = 26 + 4\cdot 10 + 26\cdot 32 = 918$.

Finally, in the case $\mathrm{Var}(\rho) > 1/16$, one may just use

$$\begin{aligned}\int_{-\infty}^{\infty}(1+x^2)\,|\varphi_\rho(x)-\varphi(x)|\,dx &\leqslant \int_{-\infty}^{\infty}(1+x^2)\varphi_\rho(x)\,dx + \int_{-\infty}^{\infty}(1+x^2)\varphi(x)\,dx \\ &= (1+\mathbf{E}\,(\rho Z)^2) + (1+\mathbf{E}\,Z^2) = 4 < 64\,\mathrm{Var}(\rho).\end{aligned}$$

Thus, Proposition 2.1 is proved. □

§3. Lower bound. Remarks on the L^p-distances

In some sense the bound of Proposition 2.1 is optimal with respect to the variance of ρ. At least, this is the case when ρ is bounded, as the following assertion shows (which is however not needed in the proof of Theorem 1.1).

Proposition 3.1. *If $\mathbf{E}\,\rho^2 = 1$ and $0 \leqslant \rho \leqslant M$ a.s., then for the distribution function Φ_ρ of the random variable $Z(\rho) = \rho Z$, where $Z \sim N(0,1)$ is independent of ρ, we have*

$$\sup_x |\Phi_\rho(x) - \Phi(x)| \geqslant \frac{c}{M^5}\,\mathrm{Var}(\rho^2), \tag{3.1}$$

where $c > 0$ is an absolute constant.

Note that the left-hand side is dominated by the total variation $\|\Phi_\rho - \Phi\|_{\mathrm{TV}}$, while $\mathrm{Var}(\rho^2) \geqslant \mathrm{Var}(\rho)$. The latter bound follows from the assumption $\rho \geqslant 0$:

$$\mathrm{Var}(\rho^2) = \mathbf{E}\,(\rho-1)^2(\rho+1)^2 \geqslant \mathbf{E}\,(\rho-1)^2 \geqslant \mathrm{Var}(\rho). \tag{3.2}$$

Proof. One may apply the following general lower bound on the Kolmogorov distance

$$\Delta = \sup_x |F_1(x) - F_2(x)|$$

between the distribution functions F_1 and F_2. Namely,

$$\Delta \geqslant \frac{1}{2\sqrt{2\pi}}\,\Big|\int_{-\infty}^{\infty} (f_1(t) - f_2(t))\,e^{-t^2/2}\,dt\Big|,$$

where f_1 and f_2 denote the characteristic functions of F_1 and F_2 respectively (cf. [10,15]). We apply it with $F_1 = \Phi_\rho$ and $F_2 = \Phi$, in which case, by Jensen's inequality,

$$f_1(t) = \mathbf{E}\,e^{-\rho^2 t^2/2} \geqslant e^{-t^2/2} = f_2(t).$$

Thus we get

$$\begin{aligned}\sup_x |\Phi_\rho(x) - \Phi(x)| &\geqslant \frac{1}{2\sqrt{2\pi}} \int_{-\infty}^{\infty} \big(\mathbf{E}\,e^{-\rho^2 t^2/2} - e^{-t^2/2}\big)\,e^{-t^2/2}\,dt \\ &= \frac{1}{2}\Big(\mathbf{E}\,\frac{1}{\sqrt{1+\rho^2}} - \frac{1}{\sqrt{2}}\Big). \end{aligned} \tag{3.3}$$

We now expand the function $u(t) = (1+t)^{-1/2}$ near the point $t_0 = 1$ according to the integral Taylor formula (2.2) up to the quadratic term. Since $u''(t) = \frac{3}{4}\,(1+t)^{-5/2}$, this gives

$$\begin{aligned}\mathbf{E}\,u(\rho^2) - u(1) &= \frac{3}{4}\,\mathbf{E}\,(\rho^2-1)^2 \int_0^1 \Big(1 + ((1-s) + s\rho^2)\Big)^{-5/2} (1-s)\,ds \\ &\geqslant \frac{3}{4}\,\mathrm{Var}(\rho^2) \int_0^1 \Big(1 + ((1-s) + sM^2)\Big)^{-5/2} (1-s)\,ds \\ &\geqslant \frac{3}{8}\,\mathrm{Var}(\rho^2)\,(1+M^2)^{-5/2},\end{aligned}$$

where we used $M \geqslant 1$ in the last step. It remains to apply (3.3) to arrive at (3.1) with $c = \frac{3}{16\cdot 2^{5/2}}$. $\square$

Proposition 2.1 may be used to obtain the (apriori weaker) non-uniform bound

$$\sup_x \Big[(1+x^2)\,|\Phi_\rho(x) - \Phi(x)|\Big] \leqslant c\,\mathrm{Var}(\rho). \tag{3.4}$$

The appearance of the weight $1 + x^2$ on the left is important in order to control the L^p-distances between Φ_ρ and Φ. Indeed, under the same assumptions as in Proposition 2.1, from (3.4) we immediately obtain:

Corollary 3.1. *For any $p \geqslant 1$,*

$$\bigg(\int_{-\infty}^{\infty} |\Phi_\rho(x) - \Phi(x)|^p\,dx\bigg)^{1/p} \leqslant c\,\mathrm{Var}(\rho),$$

where c is an absolute constant.

In particular,

$$\int_{-\infty}^{\infty} (\Phi_\rho(x) - \Phi(x))^2\,dx \leqslant c^2\,(\mathrm{Var}(\rho))^2\,.$$

In fact, here the integral on the left-hand side may easily be evaluated explicitly in terms of ρ. Indeed, let ρ' be an independent copy of ρ in order to represent the square of the characteristic function of Φ_ρ as

$$|\mathbf{E}\,e^{it\rho Z}|^2 = |\mathbf{E}\,e^{-\rho^2 t^2/2}|^2 = \mathbf{E}\,e^{-(\rho^2+\rho'^2)\,t^2/2}.$$

Hence, applying Plancherel's theorem, we get

$$\int_{-\infty}^{\infty} (\Phi_\rho(x) - \Phi(x))^2\,dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \frac{\mathbf{E}\, e^{it\rho Z} - e^{-t^2/2}}{t} \right|^2 dt$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\mathbf{E}\, e^{-(\rho^2+\rho'^2)\,t^2/2} - 2\,\mathbf{E}\, e^{-(\rho^2+1)\,t^2/2} + e^{-t^2}}{t^2}\,dt.$$

One can now apply an elementary identity

$$\int_{-\infty}^{\infty} \frac{e^{-\alpha t^2} - e^{-t^2}}{t^2}\,dt = 2\sqrt{\pi}\,\big(1 - \sqrt{\alpha}\big), \qquad \alpha \geqslant 0.$$

Indeed, the function

$$\psi(\alpha) = \int_{-\infty}^{\infty} \frac{e^{-\alpha t^2} - e^{-t^2}}{t^2}\,dt$$

is smooth on $(0,\infty)$ and has derivative $\psi'(\alpha) = -\int_{-\infty}^{\infty} e^{-\alpha t^2}\,dt = -\frac{1}{\sqrt{\alpha}}\sqrt{\pi}$. Since furthermore $\psi(1) = 0$, we get the desired assertion after integration.

Hence

$$\int_{-\infty}^{\infty} (\Phi_\rho(x) - \Phi(x))^2\,dx = \frac{1}{\sqrt{\pi}} \left[1 - 2\,\mathbf{E}\sqrt{\frac{\rho^2+1}{2}} + \mathbf{E}\sqrt{\frac{\rho^2+\rho'^2}{2}}\,\right].$$

As a result, Corollary 3.2 can be restated as follows.

Corollary 3.2. *Let $\rho \geqslant 0$ be a random variable such that $\mathbf{E}\,\rho^2 = 1$, and let ρ' be an independent copy of ρ. Then*

$$\left[1 - 2\,\mathbf{E}\sqrt{\frac{\rho^2+1}{2}} + \mathbf{E}\sqrt{\frac{\rho^2+\rho'^2}{2}}\,\right] \leqslant c\,(\mathrm{Var}(\rho))^2\,,$$

where c is an absolute constant.

It is unclear how to obtain such an estimate by a different argument (which would not be based on Proposition 2.1).

§4. Distribution of the First Coordinate on the Sphere

It remains to add the final steps in the proof of Theorem 1.1. Note that with respect to the normalized Lebesgue measure σ_{n-1} on the unit sphere S^{n-1} $(n \geqslant 2)$, the first coordinate θ_1 of a point θ is a random variable with density

$$c_n \left(1 - x^2\right)_+^{\frac{n-3}{2}}, \qquad x \in \mathbf{R}, \qquad c_n = \frac{\Gamma(\frac{n}{2})}{\sqrt{\pi}\,\Gamma(\frac{n-1}{2})},$$

where c_n is a normalizing constant. For example, when $n = 3$, this is the uniform distribution on the interval $[-1, 1]$.

Let us denote by φ_n the density of the normalized first coordinate $Z_n = \sqrt{n}\,\theta_1$ under the measure σ_{n-1}, i.e.,

$$\varphi_n(x) = c_n' \left(1 - \frac{x^2}{n}\right)_+^{\frac{n-3}{2}}, \qquad c_n' = \frac{c_n}{\sqrt{n}} = \frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi n}\;\Gamma\left(\frac{n-1}{2}\right)}.$$

The first values of these constants are $c_2' = \frac{1}{\pi\sqrt{2}} = 0.225\dots$, $c_3' = \frac{1}{2\sqrt{3}} = 0.289\dots$, $c_4' = \frac{3}{\pi} = 0.318\dots$, $c_5' = \frac{3}{4\sqrt{5}} = 0.335\dots$ Clearly, as $n \to \infty$

$$\varphi_n(x) \to \varphi(x) = \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2}, \qquad c_n' \to \frac{1}{\sqrt{2\pi}} = 0.399\dots$$

and one can show that $c_n' < \frac{1}{\sqrt{2\pi}}$ for all $n \geqslant 2$.

We are interested in non-uniform deviation bounds for $\varphi_n(x)$ from $\varphi(x)$. First let us consider the asymptotic behaviour of the functions

$$p_n(x) = \left(1 - \frac{x^2}{n}\right)_+^{\frac{n-3}{2}}, \qquad x \in \mathbf{R}.$$

Clearly, $p_n(x) \to p(x) = e^{-x^2/2}$ for all x. These functions admit a uniform Gaussian bound, since for $|x| < \sqrt{n}$ and $n \geqslant 4$,

$$-\log p_n(x) \;=\; -\frac{n-3}{2}\,\log\left(1 - \frac{x^2}{n}\right) \;\geqslant\; \frac{n-3}{2}\,\frac{x^2}{n} \;\geqslant\; \frac{x^2}{8}.$$

That is, we have:

Lemma 4.1. *If $n \geqslant 4$, then $p_n(x) \leqslant e^{-x^2/8}$ for all $x \in \mathbf{R}$.*

We also have $p_3(x) = 1$ in $|x| < \sqrt{3}$, while $p_2(x)$ is unbounded.

To study the rate of convergence of $p_n(x)$, let us derive:

Lemma 4.2. *In the interval* $|x| \leqslant \frac{1}{2}\sqrt{n}$, $n \geqslant 4$,

$$|p_n(x) - e^{-x^2/2}| \leqslant \frac{0.3}{n}\,(3x^2 + x^4)\,e^{-x^2/2}.$$

Proof. By Taylor's expansion, with some $0 \leqslant \varepsilon \leqslant 1$

$$\begin{aligned}
-\log p_n(x) &= -\frac{n-3}{2}\,\log\Big(1 - \frac{x^2}{n}\Big) \\
&= \frac{n-3}{2}\,\Big[\,\frac{x^2}{n} + \Big(\frac{x^2}{n}\Big)^2 \sum_{k=2}^{\infty} \frac{1}{k}\,\Big(\frac{x^2}{n}\Big)^{k-2}\,\Big] \\
&= \frac{n-3}{2}\,\Big(\frac{x^2}{n} + \frac{x^4}{n^2}\,\varepsilon\Big) \\
&= \frac{x^2}{2} - \frac{3x^2}{2n} + \frac{n-3}{2n^2}\,x^4\varepsilon \;=\; \frac{x^2}{2} + \frac{x^2}{2n}\,\Big(-3 + \frac{n-3}{n}\,x^2\varepsilon\Big),
\end{aligned}$$

where we assume that $|x| \leqslant \frac{1}{2}\sqrt{n}$ and $n \geqslant 4$. That is,

$$p_n(x) = p(x)\,e^{-\delta} \quad \text{with} \quad \delta = \frac{x^2}{2n}\,\Big(-3 + \frac{n-3}{n}\,x^2\varepsilon\Big).$$

Since

$$\delta \geqslant -\frac{3x^2}{2n} \geqslant -\frac{3}{8n} \geqslant -\frac{3}{32},$$

we have

$$|e^{-\delta} - 1| \leqslant |\delta|\,e^{3/32} \leqslant 1.1\,|\delta|.$$

On the other hand,

$$\delta \leqslant \frac{x^2}{2n}\,\Big(-3 + \frac{n-3}{n}\,x^2\Big) \leqslant \frac{x^4}{2n}, \qquad -\delta \leqslant \frac{3x^2}{2n},$$

which yields

$$1.1\,|\delta| \leqslant 1.1\,\Big(\frac{3x^2}{2n} + \frac{x^4}{2n}\Big) = \frac{0.55}{n}\,(3x^2 + x^4). \qquad \square$$

Combining Lemma 4.2 with Lemma 4.1, we also get a non-uniform linear bound (with respect to $1/n$) on the whole real line, namely

$$|p_n(x) - e^{-x^2/2}| \leqslant \frac{C}{n}\,e^{-x^2/16}, \qquad x \in \mathbf{R},\ n \geqslant 4,$$

where C is an absolute constant. Let us integrate this inequality over x. Since

$$\int_{-\infty}^{\infty} p_n(x)\,dx = \frac{1}{c_n'}, \qquad \int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \sqrt{2\pi},$$

we get that $|\frac{1}{c_n'} - \sqrt{2\pi}| \leqslant \frac{C}{n}$ with some absolute constant C. Using Lemmas 4.1–4.2, we then arrive at a similar conclusion about the densities φ_n.

Proposition 4.1. *If $n \geqslant 4$, then for all $x \in \mathbf{R}$, with some universal constant C*

$$|\varphi_n(x) - \varphi(x)| \leqslant \frac{C}{n}\, e^{-x^2/16}. \tag{4.1}$$

Proof of Theorem 1.1. Assuming (without loss of generality) that $n \geqslant 4$, let Φ_n and φ_n denote respectively the distribution function and density of $Z_n = \theta_1 \sqrt{n}$, where θ_1 is the first coordinate of a random point uniformly distributed in the unit sphere S^{n-1}. If $\rho^2 = \frac{1}{n}\,|X|^2$ is independent of Z_n ($\rho \geqslant 0$), then, by the definition of the typical distribution,

$$F(x) = \mathbf{P}\{\rho Z_n \leqslant x\} \,=\, \mathbf{E}\,\Phi_n(x/\rho), \qquad x \in \mathbf{R},$$

so that

$$\int_{-\infty}^{\infty} (1+x^2)\,|F(dx) - \Phi_\rho(dx)| \,=\, \int_{-\infty}^{\infty} (1+x^2)\,|\mathbf{E}\,\Phi_n(dx/\rho) - \mathbf{E}\,\Phi(dx/\rho)|. \tag{4.2}$$

But, for any fixed value of ρ,

$$\int_{-\infty}^{\infty} (1+x^2)\,|\Phi_n(dx/\rho) - \Phi(dx/\rho)|(dx) \,=\, \int_{-\infty}^{\infty} (1+\rho^2 x^2)\,|\Phi_n(dx) - \Phi(dx)|,$$

so that, by (4.2), taking the expectation with respect to ρ and using Jensen's inequality, we get

$$\begin{aligned}\int_{-\infty}^{\infty} (1+x^2)\,|F(dx) - \Phi_\rho(dx)| \;&\leqslant\; \mathbf{E}\int_{-\infty}^{\infty} (1+x^2)\,|\Phi_n(dx/\rho) - \Phi(dx/\rho)| \\ &=\; \mathbf{E}\int_{-\infty}^{\infty} (1+\rho^2x^2)\,|\Phi_n(dx) - \Phi(dx)| \\ &=\; \int_{-\infty}^{\infty} (1+x^2)\,|\Phi_n(dx) - \Phi(dx)|.\end{aligned}$$

It remains to apply (4.1), which yields

$$\int_{-\infty}^{\infty} (1+x^2)\,|\Phi_n(dx) - \Phi(dx)| \,=\, \int_{-\infty}^{\infty} (1+x^2)\,|\varphi_n(x) - \varphi(x)|\,dx \,\leqslant\, \frac{C}{n}$$

with some universal constant C. □

References

1. M. Anttila, K. Ball, I. Perissinaki, *The central limit problem for convex bodies.* — Trans. Amer. Math. Soc. **355**, No. 12 (2003), 4723–4735.
2. S. G. Bobkov, *On concentration of distributions of random weighted sums.* — Ann. Probab. **31**, No. 1 (2003), 195–215.
3. S. G. Bobkov, *Concentration of distributions of the weighted sums with Bernoullian coefficients.* — Geometric aspects of functional analysis, Lecture Notes in Math., **1807**, pp. 27–36, Springer, Berlin, 2003.
4. S. G. Bobkov, *Concentration of normalized sums and a central limit theorem for noncorrelated random variables.* — Ann. Probab. **32**, No. 4 (2004), 2884–2907.
5. S. G. Bobkov, *On a theorem of V. N. Sudakov on typical distributions.* — J. Math. Sciences (New York), **167**, No. 4 (2010), 464–473. Translated from: Zap. Nauchn. Semin. POMI, **368** (2009), 59–74 (Russian).
6. S. G. Bobkov, F. Götze, *On the central limit theorem along subsequences of noncorrelated observations.* — Teor. Veroyatnost. i Primenen. 48, No. 4 (2003), 745–765 (Russian). Reprinted in: Theory Probab. Appl. **48**, No. 4 (2004), 604–621.
7. S. G. Bobkov, F. Götze, *Concentration inequalities and limit theorems for randomized sums.* — Probab. Theory Rel. Fields, **137**, No. 1–2 (2007), 49–81.
8. S. G. Bobkov, A. Koldobsky, *On the central limit property of convex bodies.* — Geometric aspects of functional analysis, Lecture Notes in Math. **1807**, pp. 44–52, Springer, Berlin, 2003.
9. R. Eldan, B. Klartag, *Pointwise estimates for marginals of convex bodies.* — J. Funct. Anal. **254**, No. 8 (2008), 2275–2293.
10. I. A. Ibragimov, Yu. V. Linnik, *Independent and Stationary Sequences of Random Variables.* With a supplementary chapter by I. A. Ibragimov and V. V. Petrov. Translation from the Russian edited by J. F. C. Kingman. Wolters-Noordhoff Publishing, Groningen, 1971, 443 pp.
11. B. Klartag, *A central limit theorem for convex sets.* — Invent. Math. **168**, No. 1 (2007), 91–131.
12. B. Klartag, *Power-law estimates for the central limit theorem for convex sets.* — J. Funct. Anal. **245**, No. 1 (2007), 284–310.
13. B. Klartag, *Uniform almost sub-Gaussian estimates for linear functionals on convex sets.* — Algebra i Analiz **19**, No. 1 (2007), 109–148; reprinted in St. Petersburg Math. J. **19**, No. 1 (2008), 77–106.
14. B. Klartag, *A Berry–Esseen type inequality for convex bodies with an unconditional basis.* — Probab. Theory Rel. Fields **145**, No. 1–2 (2009), 1–33.

15. V. K. Matskyavichyus, *A lower bound for the rate of convergence in the central limit theorem.* — Teor. Veroyatnost. i Primenen. **28**, No. 3 (1983), 565–569 (Russian).
16. M. W. Meckes, *Gaussian marginals of convex bodies with symmetries.* — Beiträge Algebra Geom. **50**, No. 1 (2009), 101–118.
17. S. V. Nagaev, *On the distribution of linear functionals in finite-dimensional spaces of large dimension.* — Dokl. Akad. Nauk SSSR **263**, No. 2 (1982), 295–297 (Russian).
18. S. Sodin, *Tail-sensitive Gaussian asymptotics for marginals of concentrated measures in high dimension.* — Geometric aspects of functional analysis. Lecture Notes in Math. **1910**, pp. 271–295, Springer, Berlin, 2007.
19. A. V. Sudakov, V. N. Sudakov, *Empirical distributions and dimension-invariant estimates for spread in the problem of typical distributions.* — Zap. Nauchn. Semin. POMI **260** (1999), 258–262 (Russian); translation in J. Math. Sci. (New York) **109**, No. 6 (2002), 2225–2228.
20. V. N. Sudakov, *Typical distributions of linear functionals in finite-dimensional spaces of high dimension.* — Soviet Math. Dokl. **19** (1978), 1578–1582 (Russian); translation in: Dokl. Akad. Nauk SSSR, **243**, No. 6 (1978), 1402–1405.
21. V. N. Sudakov, *Distributions of randomly selected random variables.* — Zap. Nauchn. Semin. POMI **216** (1994), 153–160; translation in J. Math. Sci. (New York) **88**, No. 1 (1998), 106–109.
22. H. von Weizsäcker, *Sudakov's typical marginals, random linear functionals and a conditional central limit theorem.* — Probab. Theory Rel. Fields, **107** (1997), 313–324.

School of Mathematics,
University of Minnesota 127 Vincent Hall,
206 Church St. S.E., Minneapolis,
MN 55455 USA
E-mail: bobkov@math.umn.edu

Поступило 13 марта 2017 г.

Fakultät für Mathematik,
Universität Bielefeld Postfach
100131, 33501 Bielefeld, Germany
E-mail: chistyak@math.uni-bielefeld.de

Faculty of Mathematics, University of Bielefeld, Germany
E-mail: goetze@math.uni-bielefeld.de