J. Math. Anal. Appl. 354 (2009) 704–707

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

www.elsevier.com/locate/jmaa

A note on multivariate Gaussian estimates

Dawei Lu [a,*], Wenbo V. Li [b,1]

[a] Department of Applied Mathematics, Dalian University of Technology, Dalian 116023, China
[b] Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, United States

ARTICLE INFO

Article history:
Received 13 November 2008
Submitted by M. Peligrad

Keywords:
Multivariate Mills' ratio

ABSTRACT

An upper bound of multivariate Gaussian probability for a general convex domain D is given based on a geometric observation. The bound is sharper than known ones on multivariate Mills' ratio in many cases.

1. Introduction

Bounds on multivariate Gaussian probability over a suitable region arise naturally in many areas of probability and statistics. There is also a long history of studying them in various settings, in particular the case of multivariate Mills' ratio given in (3.1). Two well-known approaches are employed in the literature, see for example Hashorva and Hüsler [3]. One is translating multivariate Mills' ratio into a product of univariate Mills' ratios and then using corresponding lower and upper bounds of univariate Mills' ratio to estimate each term. The key point of this approach is based on the maximal and minimal eigenvalues of the covariance matrix which are hard to find. The second approach is using Jensen's inequality to find a lower bound. The basic idea is separating out diagonal terms in the Gaussian density and then using the fact that remaining density term is a convex function.

In this note, we provide a new upper bound approach for general convex domain based on a geometric observation associate with a dominating point under the given Gaussian density. Our bound is sharper than known ones on multivariate Mills' ratio in many case. In order to do the comparison, we also present some useful facts on univariate Mills' ratio.

The rest of this paper is arranged as follows. In Section 2, we write out explicit coefficients of the polynomial in the numerator of the nth approximation of Mills' ratio. Two related functions and their limits are also given. In Section 3, we briefly review known approaches of multivariate Mills' ratio and give our main result of this paper. An example in Hashorva and Hüsler [3] is analyzed and show that for certain regions of parameters, our upper bound is better than those listed in their paper.

2. Univariate Mills' ratio

Mills' ratio is defined as

$$R(x) = e^{x^2/2} \int_x^\infty e^{-t^2/2}\,dt$$

* Corresponding author.
E-mail addresses: ludawei_dlut@163.com (D. Lu), wli@math.udel.edu (W.V. Li).
1 Supported by NSF grant DMS-0805929.

0022-247X/$ – see front matter
doi:10.1016/j.jmaa.2009.01.046

which can be approximated by

$$R(x) = 1/x - 1/x^3 + 1 \cdot 3/x^5 - \cdots = \sum_{n=0}^{\infty} (-1)^n \frac{(2n)!}{2^n n!} x^{-2n-1}.$$

The well-known bounds, which can be checked by differentiation, are

$$x/\left(x^2+1\right) \leqslant R(x) \leqslant 1/x, \quad x > 0. \tag{2.1}$$

In fact, it is easy to see from Shenton [8]

$$\frac{x}{x^2+1} = R_2(x) \leqslant R_4(x) \leqslant \cdots \leqslant R(x) \leqslant \cdots \leqslant R_3(x) \leqslant R_1(x) = \frac{1}{x}, \quad x > 0, \tag{2.2}$$

where $R_n(x)$ is the nth approximation of the continued fraction

$$R(x) = \frac{1}{x+}\,\frac{1}{x+}\,\frac{2}{x+}\,\frac{3}{x+}\cdots.$$

It is also known that

$$R_n(x) = p_n(x)/q_n(x),$$

where

$$q_n(x) = \sum_{0 \leqslant 2m \leqslant n} \frac{(2m)!}{2^m m!} \binom{n}{2m} x^{n-2m}$$

is the Hermite polynomial of degree n with leading coefficient 1, and

$$p_n(x) = \sum_{0 \leqslant i \leqslant n-1} q_{n-1-i}^{(i)}(x) \tag{2.3}$$

is a polynomial of degree $n-1$ with leading coefficient 1. Here $q_n^{(k)}(x)$ is the kth derivative with respect to x. The monotonicity of error of the approximation is discussed in Dudley [2].

Here, we provide a concise expression for $p_n(x)$ which is useful in Section 3.

Proposition 2.1. *We can write*

$$p_n(x) = \sum_{0 \leqslant 2m \leqslant n-1} \left(\frac{(n-1-m)!}{2^m (n-1-2m)!} \sum_{0 \leqslant i \leqslant m} \binom{n}{i} \right) x^{n-1-2m}.$$

Proof. The aim is to rewrite (2.3) into a polynomial in x with explicit coefficients. Note that $q_n(x) = E(x+\xi)^n$, where ξ is a standard normal random variable. From (2.3) we have

$$p_n(x) = \sum_{0 \leqslant i \leqslant n-1} \left(E(x+\xi)^{n-1-i}\right)^{(i)} = \sum_{0 \leqslant 2i \leqslant n-1} \frac{(n-1-i)!}{(n-1-2i)!} E(x+\xi)^{n-1-2i}.$$

Using the fact $E\xi^{2k} = (2k)!/2^k k!$ and $E\xi^{2k+1} = 0$,

$$E(x+\xi)^{n-1-2i} = \sum_{0 \leqslant 2k \leqslant n-1-2i} \binom{n-1-2i}{2k} \frac{(2k)!}{2^k k!} x^{n-1-2i-2k}.$$

Combining them together, and replacing $i+k$ by m,

$$p_n(x) = \sum_{0 \leqslant 2i \leqslant n-1} \sum_{2i \leqslant 2m \leqslant n-1} \frac{(n-1-i)!}{(n-1-2m)!} \frac{1}{2^{m-i}(m-i)!} x^{n-1-2m}.$$

Exchanging the summations and replacing $m-i$ by j, we have

$$p_n(x) = \sum_{0 \leqslant 2m \leqslant n-1} \sum_{0 \leqslant j \leqslant m} \frac{(n-1-m+j)!}{(n-1-2m)!} \frac{1}{2^j j!} x^{n-1-2m}.$$

The crucial observation is the identity

$$\sum_{0\leqslant j\leqslant m}\frac{1}{2^j}\binom{n-1-m+j}{n-1-m}=\frac{1}{2^m}\sum_{0\leqslant i\leqslant m}\binom{n}{i},$$

which can be checked by induction on m. It can also be proved directly by a technique detailed in Li [5]. Thus a simple substitution finishes the proof. □

Next we discuss some better bounds of $R(x)$ than those in (2.1). It was given in Birnbaum [1] and Sampford [6], that

$$2/\big(\sqrt{x^2+4}+x\big)<R(x)<4/\big(\sqrt{x^2+8}+3x\big),\quad x>0.$$

Thus one can consider a function $\theta(x)$ such that

$$R(x)=4/\big(\sqrt{x^2+\theta(x)}+3x\big),\quad x>0. \tag{2.4}$$

Here are some basic properties of $\theta(x)$.

Proposition 2.2. *The function $\theta(x)$ is well-defined on $(0,\infty)$, and*

$$\theta_1(x)\leqslant\theta_3(x)\leqslant\cdots\leqslant\theta(x)\leqslant\cdots\leqslant\theta_4(x)\leqslant\theta_2(x),\qquad \lim_{x\to\infty}\theta(x)=8, \tag{2.5}$$

where $\theta_n(x)=8(1/R_n(x)-x)(2/R_n(x)-x)$.

Proof. From (2.4), we can rewrite $\theta(x)=8(1/R(x)-x)(2/R(x)-x)$. For $x>0$ fixed and $0<y<4/3x$, it is easy to see that $f(y)=8(1/y-x)(2/y-x)$ is decreasing. Hence the bounds in (2.5) follow from (2.2). Furthermore, $\theta_2(x)=8(x^2+2)/x^2$ and $\theta_3(x)=8(x^4+4x^2)/(x^4+4x^2+4)$, and the common limit is 8, which implies $\lim_{x\to\infty}\theta(x)=8$. We finish the whole proof. □

We can also consider a function $\beta(x)$ such that $R(x)=2/(x+\sqrt{x^2+\beta(x)})$, and $\beta_n(x)=4(1/R_n(x)-x)/R_n(x)$. Then $\beta(x)$ is well-defined on $(0,\infty)$, and

$$\beta_1(x)\leqslant\beta_3(x)\leqslant\cdots\leqslant\beta(x)\leqslant\cdots\leqslant\beta_4(x)\leqslant\beta_2(x),\qquad \lim_{x\to\infty}\beta(x)=4.$$

The proof is similar and we omit the details.

3. Results on general convex domain

We first briefly review some known results and methods used in particular for the convex domain $D=\{X\geqslant a\}$, where $X=(X_1,\dots,X_n)^T$ and $a=(a_1,\dots,a_n)^T$, $\{X\geqslant a\}=\bigcap_{i=1}^n\{X_i\geqslant a_i\}$. Let X be a multivariate normal random vector with positive definite covariance matrix Σ and $M=\Sigma^{-1}$. The multivariate Mills' ratio $R(a,M)$ is defined by

$$R(a,M)=(2\pi)^{n/2}|\Sigma|^{1/2}\exp\big(a^TMa/2\big)\cdot P(X\geqslant a). \tag{3.1}$$

Note that $R(a,M)$ is the multivariate normal probability beyond a certain point divided by the multivariate normal density at that point. It is easy to obtain by shift substitution

$$R(a,M)=\int\limits_{u\geqslant 0}\exp\big(-a^TMu-u^TMu/2\big)\,du,$$

where $u=(u_1,\dots,u_n)^T$.

In Savage [7], a pair of upper and lower bounds for $R(a,M)$ was given by using $1-x\leqslant e^{-x}\leqslant 1$ on $\exp(-u^TMu/2)$. In Steck [9], two lower bounds were given by Jensen's inequality as outline in the introduction. In Hashorva and Hüsler [3], a pair of upper and lower bounds for multivariate Gaussian tails probability was given by the minimal and maximal eigenvalues after a shift substitution. They also obtained another lower bound by Jensen's inequality again after a shift substitution.

Below is an upper bound of multivariate Gaussian probability for a general convex domain based on a geometric idea used in Kuelbs, Li and Linde [4] in a special setting. The shift substitution using the dominating point plays a crucial role.

Theorem 3.1. *Let X be a centered multivariate normal random vector with positive definite covariance matrix Σ, and $M=\Sigma^{-1}$. Let D be a closed convex domain which does not contain the origin, and x^* be the unique closest point in D to the origin under Hilbert norm $\langle x,Mx\rangle^{1/2}$. Then*

$$P(X\in D)\leqslant\exp\big(-\langle x^*,Mx^*\rangle/2\big)\cdot P(X\in D-x^*). \tag{3.2}$$

In particular, $P(X\in D)\leqslant(1/2)\cdot\exp(-\langle x^*,Mx^*\rangle/2)$.

Proof. First note that $\langle x, My\rangle$ is an inner product on R^n. Indeed, the associated Hilbert space is the reproducing kernel Hilbert space generated by the Gaussian measure $\mu(A) = P(X \in A)$. Thus $x^* \in D$ is uniquely defined, see [4].

By the convexity of D and definition of x^*, for arbitrary $x \in D$ and $0 \leqslant \lambda \leqslant 1$, $\lambda x^* + (1-\lambda)x \in D$ and

$$\big\langle \lambda x^* + (1-\lambda)x, M\big(\lambda x^* + (1-\lambda)x\big)\big\rangle \geqslant \langle x^*, Mx^*\rangle. \tag{3.3}$$

Simplify (3.3) and cancel the factor $(1-\lambda)$, we obtain

$$2\lambda\langle x, Mx^*\rangle + (1-\lambda)\langle x, Mx\rangle \geqslant (1+\lambda)\langle x^*, Mx^*\rangle.$$

Taking $\lambda \to 1$, we have for all $x \in D$,

$$\langle x, Mx^*\rangle \geqslant \langle x^*, Mx^*\rangle. \tag{3.4}$$

Next, by shift substitution $y = x - x^*$, and using (3.4),

$$\begin{aligned} P(X \in D) &= \frac{|M|^{1/2}}{(2\pi)^{n/2}} \int\limits_{y \in D - x^*} \exp\big(-\big(\langle y, My\rangle + 2\langle y, Mx^*\rangle + \langle x^*, Mx^*\rangle\big)/2\big)\,dy \\ &\leqslant \frac{|M|^{1/2}}{(2\pi)^{n/2}} \int\limits_{y \in D - x^*} \exp\big(-\big(\langle y, My\rangle + \langle x^*, Mx^*\rangle\big)/2\big)\,dy \\ &= \exp\big(-\langle x^*, Mx^*\rangle/2\big) \cdot P(X \in D - x^*). \end{aligned}$$

Finally note that $D - x^* \subset \{x\colon \ \langle x, Mx^*\rangle \geqslant 0\}$ which is a half-space with Gaussian measure $1/2$. We finish the whole proof. □

Next, we consider a simple example which was examined in detail in [3]. We show that the upper bound in (3.2) is better than those listed in [3] for certain range of parameters. Let ξ, η be two Gaussian random variables with mean zero and the covariance matrix $\Sigma = \binom{1\ \rho}{\rho\ 1}$, $-1 < \rho < 1$, and $M = \Sigma^{-1}$. Fix $D = \{x_1 \geqslant b_1, x_2 \geqslant b_2\}$ with constants $b_1 > b_2 > 0$. In the case $0 < \rho < b_2/b_1$, the upper bound given in [3] is equivalent to

$$P(\xi \geqslant b_1, \eta \geqslant b_2) \leqslant \exp\big(-\langle x^*, Mx^*\rangle/2\big) \cdot \frac{1+\rho}{2\pi(1-\rho^2)^{1/2}} R(z_1)R(z_2), \tag{3.5}$$

where $z_1 = (b_1 - \rho b_2)(1+\rho)^{1/2}/(1-\rho^2)$, $z_2 = (b_2 - \rho b_1)(1+\rho)^{1/2}/(1-\rho^2)$. For the bound from Theorem 3.1, we have

$$\begin{aligned} P(\xi \geqslant b_1, \eta \geqslant b_2) &\leqslant \exp\big(-\langle x^*, Mx^*\rangle/2\big) \cdot P(\xi \geqslant 0, \eta \geqslant 0) \\ &= \exp\big(-\langle x^*, Mx^*\rangle/2\big) \cdot \frac{\pi/2 + \arcsin(\rho)}{2\pi}. \end{aligned} \tag{3.6}$$

Note that for $0 < \rho < b_2/b_1$, we have

$$(1-\rho)(b_1 + b_2) > b_1 - \rho b_2 > b_1(1-\rho) > b_2(1-\rho) > b_2 - \rho b_1 > 0.$$

Since $R(x)$ is decreasing on $(0, \infty)$, we see that

$$R(z_1)R(z_2) > R\big((b_1 + b_2)/(1+\rho)^{1/2}\big) \cdot R\big(b_2/(1+\rho)^{1/2}\big) > R(2b_1)R(b_1). \tag{3.7}$$

Using (2.2), (3.7) and the fact $\arcsin(\rho) \leqslant \rho\pi/2$ on $[0, 1]$, a sufficient condition of the upper bound in (3.6) is less than the one in (3.5) is

$$b_2/b_1 > \rho > \big(1 - 4R_{2n}^2(2b_1) \cdot R_{2n}^2(b_1)/\pi^2\big)^{1/2}, \quad \text{for any } n \geqslant 1.$$

The range for ρ above is non-empty. In fact for $n = 1$, one can take $1 > b_2/b_1 > \sqrt{1 - 4/25\pi^2}$ and $1 > b_1 > (5\pi/2)^{1/2} \times (1 - b_2^2/b_1^2)^{1/4}$.

References

[1] Z.W. Birnbaum, An inequality for Mill's ratio, Ann. Math. Statist. 13 (1942) 245–246.
[2] R.M. Dudley, Some inequalities for continued fraction, Math. Comp. 49 (1987) 585–593.
[3] E. Hashorva, J. Hüsler, On multivariate Gaussian tails, Ann. Inst. Statist. Math. 55 (3) (2003) 507–522.
[4] J. Kuelbs, W.V. Li, W. Linde, The Gaussian measure of shifted balls, Probab. Theory Related Fields 98 (1994) 143–162.
[5] W.V. Li, Discover binomial coefficient identities via calculus, preprint, 2008.
[6] M.R. Sampford, Some inequalities on Mill's ratio and related functions, Ann. Math. Statist. 24 (1953) 130–132.
[7] I.R. Savage, Mills' ratio for multivariate normal distributions, J. Res. Nat. Bur. Stand. Sect. B 66 (1962) 93–96.
[8] L.R. Shenton, Inequalities for the normal integral including a new continued fraction, Biometrika 41 (1954) 177–189.
[9] G.P. Steck, Lower bounds for the multivariate normal Mills' ratio, Ann. Probab. 7 (1979) 547–551.