JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **102**, 149–155 (1984)

# An Inequality for $L_p$-Norms with Respect to the Multivariate Normal Distribution

W. STADJE

*University of Göttingen*

*Submitted by R. P. Boas*

A partial converse of Jensen's inequality for integrals of norms on $\mathbb{R}^k$ is proved.

Let $\|\cdot\|$ be an arbitrary norm on $\mathbb{R}^k$, $A$ a positive definite $(k \times k)$-matrix and $N_A$ the normal distribution on $\mathbb{R}^k$ with mean 0 and covariance matrix $A$. The purpose of this note is the proof of the following:

THEOREM. *For* $p \in [1, \infty)$

$$\left(\int \|x\|^p \, N_A(dx)\right)^{1/p} \leqslant 2\Gamma\left(\frac{p+1}{2}\right)^{1/p} \pi^{(p-1)/2p} \int \|x\| \, N_A(dx). \tag{1}$$

If we denote the $L_p$-norm of a measurable function $f\colon \mathbb{R}^k \to \mathbb{R}$ with respect to $N_A$ by $\|f\|_{p,N_A}$, the theorem says that

$$\| \|\cdot\| \|_{p,N_A} \leqslant c_p \| \|\cdot\| \|_{1,N_A}, \tag{1'}$$

where $c_p$ is a constant only depending on $p$, especially not on the norm $\|\cdot\|: c_p := 2\Gamma((p+1)/2)^{1/p} \pi^{(p-1)/2p}$. It is a simple consequence of Jensen's inequality that $\|f\|_{1,N_A} \leqslant \|f\|_{p,N_A}$ for $p \geqslant 1$. Thus (1′) can be considered as a partial converse of this inequality for a special family of functions [1, p. 39]. If we take for example $\|x\| := (\sum_{i=1}^k \alpha_i |x_i|^p)^{1/p}$, where $\alpha_1, ..., \alpha_k > 0$, (1′) transforms into

$$\int \left(\sum_{i=1}^k \alpha_i |x_i|^p\right)^{1/p} N_A(dx) \geqslant \frac{1}{c_p} \left(\sum_{i=1}^k \alpha_i \int |x_i|^p \, N_A(dx)\right)^{1/p}. \tag{2}$$

The right-hand side of (2) can be computed directly, and if the variances of the components $x_1, ..., x_n$ (i.e., the diagonal elements of $A$) are denoted by $\sigma_1^2, ..., \sigma_n^2$, we get

$$\int \left(\sum_{i=1}^k \alpha_i |x_i|^p\right)^{1/p} N_A(dx) \geqslant (2\pi)^{-1/2} \left(\sum_{i=1}^k \alpha_i \sigma_i^p\right)^{1/p}, \tag{3}$$



0022-247X/84 $3.00

which in the case of the standard normal distribution becomes

$$\int \left(\sum_{i=1}^{k} \alpha_i |x_i|^p\right)^{1/p} e^{-|x|^2/2}\, \lambda^k(dx) \geqslant (2\pi)^{(k-1)/2} \left(\sum_{i=1}^{k} \alpha_i\right)^{1/p}. \tag{4}$$

The proof of (1) depends on the following lemma whose easy proof is included for completeness. A function $f\colon (a,b) \to [0,\infty)$ satisfying $f \geqslant 0$ and $\int_a^b f(t)\,dt = 1$ is called a density function on $(a,b)$. Further we set $\|f\|_p := (\int_a^b f^p\,dt)^{1/p}$.

LEMMA. *Let $f$ and $g$ be two positive functions on the interval $(a,b)$ such that $f, g$ and $g/f$ are monotone nondecreasing. Then we have for all $p \in [1,\infty]$*

$$\|f\|_p \leqslant \|g\|_p. \tag{5}$$

*If equality holds for some $p \in (1,\infty)$, it follows that $f = g$ Lebesgue-almost everywhere.*

*Proof.* Since the case $p \in \{1,\infty\}$ is trivial, let $p \in (1,\infty)$. There are numbers $a \leqslant c \leqslant d \leqslant b$ satisfying $\{t \in (a,b) \mid g(t) < f(t)\} \subset (a,c]$, $\{t \in (a,b) \mid g(t) > f(t)\} \subset [d,b)$, $\{t \in (a,b) \mid g(t) = f(t)\} \supset (c,d)$. For $0 \leqslant \lambda \leqslant 1$ and $h_\lambda := p(\lambda f + (1-\lambda)\,g)^{p-1}$ we clearly have $\sup_{(a,c]} h_\lambda \leqslant \sup_{[d,b)} h_\lambda$. If $\|f\|_p = \infty$, obviously

$$\int_d^b g^p\,dt \geqslant \int_d^b f^p\,dt = \infty$$

implying (5). Thus, without loss of generality we may assume $\|f\|_p < \infty$, $\|g\|_p < \infty$. Then we obtain using $\int_a^c (f-g)\,dt = -\int_d^b (f-g)\,dt$:

$$\begin{aligned}
(d/d\lambda) &\left\{\int_a^b [\lambda f + (1-\lambda)\,g]^p\,dt\right\} \\
&= p\int_a^b [\lambda f + (1-\lambda)\,g]^{p-1}(f-g)\,dt \\
&\leqslant \sup_{(a,c)} h_\lambda \int_a^c (f-g)\,dt + \inf_{(d,b)} h_\lambda \int_d^b (f-g)\,dt \\
&= \left(\inf_{(d,b)} h_\lambda - \sup_{(a,c)} h_\lambda\right) \int_d^b (f-g)\,dt \leqslant 0.
\end{aligned} \tag{6}$$

Hence $\lambda \to \|\lambda f + (1-\lambda)\,g\|_p$ decreases monotonically on $[0,1]$ and attains its minimum for $\lambda = 1$ and its maximum for $\lambda = 0$. This proves (5). Now let

$f \neq g$ with respect to Lebesgue measure so that $\int_d^b (f-g)\,dt < 0 < \int_a^c (f-g)\,dt$. $\lambda \to \|\lambda f + (1-\lambda) g\|_p$ can be a constant function only if for all $\lambda \in [0, 1]$ $h_\lambda = \text{const.}$ on $(a, c)$ and on $(d, b)$. As $g/f$ is monotone nondecreasing, in that case both $f$ and $g$ have to be constant on $(a, c)$ as well as on $(d, b)$: $f \mid (a, c) = \alpha_1 > \alpha_2 = g \mid (a, c)$, $f \mid (d, b) = \beta_1 < \beta_2 = g \mid (d, b)$. Since the derivative in (6) vanishes, we must have $\inf_{(d,b)} h_\lambda = \sup_{(a,c)} h_\lambda$, and this implies $\lambda\alpha_1 + (1-\lambda)\alpha_2 = \lambda\beta_1 + (1-\lambda)\beta_2$ for all $\lambda \in [0, 1]$. Thus we can conclude $f = g = 1$ contradicting the assumption $f \neq g$.

*Proof of the Theorem.* It is sufficient to show (1) for the standard normal distribution $N$, as is easily seen by the substitution $y = A^{-1/2}x$, for $y \to \|A^{1/2}y\|$ is again a norm. Let $F(t) := N\{x \mid \|x\| \leqslant t\}$ be the distribution function of $\|\cdot\|$ with respect to $N$ and $\Phi(t) := (2\pi)^{-1/2} \int_{-\infty}^t \exp\{-u^2/2\}\,du$ be the standard normal distribution function. Then we have the relations

$$\int \|x\|\, N(dx) \geqslant \int_{1/2}^1 F^{-1}(t)\,dt, \tag{7}$$

$$\int \|x\|^p\, N(dx) = \int_0^1 |F^{-1}(t)|^p\,dt \leqslant 2\int_{1/2}^1 F^{-1}(t)^p\,dt, \tag{8}$$

$$\begin{aligned} 2\int_{1/2}^1 \Phi^{-1}(t)^p\,dt &= \int_0^1 |\Phi^{-1}(t)|^p\,dt = \int_{-\infty}^{\infty} |t|^p\,d\Phi(t) \\ &= (2\pi)^{-1/2}\, 2^{(p+1)/2}\, \Gamma\left(\frac{p+1}{2}\right) = \pi^{-1/2}\, 2^{p/2}\, \Gamma\left(\frac{p+1}{2}\right). \end{aligned} \tag{9}$$

$$\frac{F^{-1}}{\int_{1/2}^1 F^{-1}(t)\,dt} \quad \text{and} \quad \frac{\Phi^{-1}}{\int_{1/2}^1 \Phi^{-1}(t)\,dt}$$

are monotone nondecreasing density functions on $(\frac{1}{2}, 1)$. Below we shall show that $\Phi^{-1}/F^{-1}$ is also monotone nondecreasing. Using the lemma we can therefore conclude from (7)–(9),

$$\begin{aligned} \int \|x\|^p\, N(dx) &\leqslant 2\int_{1/2}^1 F^{-1}(t)^p\,dt \\ &\leqslant 2\,\frac{[\int_{1/2}^1 F^{-1}(t)\,dt]^p}{[\int_{1/2}^1 \Phi^{-1}(t)\,dt]^p} \int_{1/2}^1 \Phi^{-1}(t)^p\,dt \\ &\leqslant \frac{[\int \|x\|\, N(dx)]^p}{[\pi^{-1/2}2^{-1/2}]^p}\, \pi^{-1/2}\, \Gamma\left(\frac{p+1}{2}\right) 2^{p/2} \\ &= \pi^{(p-1)/2}\, \Gamma\left(\frac{p+1}{2}\right) 2^p \left[\int \|x\|\, N(dx)\right]^p. \end{aligned} \tag{10}$$

Thus it remains to prove that $\Phi^{-1}/F^{-1}$ is monotone nondecreasing. $F$ is differentiable on $(0, \infty)$, because we have for $t > 0$ and $\alpha_\varepsilon := (t+\varepsilon)/t$:

$$\lim_{\varepsilon \to 0} \varepsilon^{-1}(F(t+\varepsilon) - F(t)) = \lim_{\varepsilon \to 0} (t(\alpha_\varepsilon - 1))^{-1}(F(\alpha_\varepsilon t) - F(t))$$
$$= t^{-1} u_t'(1), \tag{11}$$

where $u_t(\alpha) := F(\alpha t)$, $\alpha > 0$. $u_t$ is differentiable because of

$$(d/d\alpha)\, u_t(\alpha) = (d/d\alpha) \int_{\{x \mid \|x\| \leqslant \alpha t\}} (2\pi)^{-k/2} \exp\left\{ -\frac{|x|^2}{2} \right\} \lambda^k(dx)$$
$$= (d/d\alpha) \int_{\{v \mid \|v\| \leqslant t\}} \alpha^k \exp\{-(1/2)\,\alpha^2\,|v|^2\}(2\pi)^{-k/2}\lambda^k(dv)$$
$$= (2\pi)^{-k/2}\,\alpha^{k-1}\,k \int_{\{v \mid \|v\| \leqslant t\}} \exp\{-(1/2)\,\alpha^2\,|v|^2\}\,\lambda^k(dv)$$
$$- (2\pi)^{-k/2}\,\alpha^{k+1} \int_{\{v \mid \|v\| \leqslant t\}} |v|^2$$
$$\times \exp\{-(1/2)\,\alpha^2\,|v|^2\}\,\lambda^k(dv). \tag{12}$$

By (11) and (12) we get

$$F'(t) = t^{-1}\left[ kF(t) - \int_{\{v \mid \|v\| \leqslant t\}} |v|^2 (2\pi)^{-k/2} \exp\{-(1/2)\,|v|^2\}\,\lambda^k(dv) \right]. \tag{13}$$

Similarly it is seen that

$$\Phi'(s) = s^{-1}\left( k\Phi(s) - \int_{\{v \mid v_1 \leqslant s\}} |v|^2 (2\pi)^{-k/2} \exp\left\{ -\frac{|v|^2}{2} \right\} \lambda^k(dv), \tag{14}$$

where $v = (v_1, \ldots, v_k)$. Now let

$$r = \Phi(s) = F(t) \text{ for some } s, t > 0 \text{ and } r \in (1/2, 1). \tag{15}$$

We denote by $\mu_{k,r}$ the normalized surface measure on the sphere $rS^{k-1}$. By a theorem due to Landau and Shepp ([4, p. 373, Lemma 1] (condition (3.8) there is superfluous)), for each half space $H$ containing the origin as an interior point and each closed convex set $C \subset \mathbb{R}^k$ the function $r \to \mu_{k,r}(H) - \mu_{k,r}(C)$ changes its sign at most once, and if it does, it is positive for large $r$ (see also the remark after this proof). If we take

$H=\{v\in\mathbb{R}^k \mid v_1\leqslant s\}$ and $C=\{v\in\mathbb{R}^k \mid \|v\|\leqslant t\}$, we obtain for $r_0=\inf\{r>0 \mid \mu_{k,r}(H)>\mu_{k,r}(C)\}$:

$$
\begin{aligned}
&\int_{\mathbb{R}^k} |v|^2\, e^{-(1/2)|v|^2}(1_H(v)-1_C(v))\,\lambda^k(dv)\\
&\quad=\frac{2\pi^{k/2}}{\Gamma(k/2)}\int_0^\infty r^{k+1}(\mu_{k,r}(H)-\mu_{k,r}(C))\,e^{-r^2/2}\,dr\\
&\quad=\frac{2\pi^{k/2}}{\Gamma(k/2)}\int_0^\infty (r^2-r_0^2)\,r^{k-1}(\mu_{k,r}(H)-\mu_{k,r}(C))\;e^{-r^2/2}\,dr\\
&\quad\geqslant 0.
\end{aligned}
\tag{16}
$$

The first equation follows from introducing polar coordinates and the obvious identity

$$
\begin{aligned}
\mu_{k,r}(E)=\frac{\Gamma(k/2)}{2\pi^{k/2}}\int_0^{2\pi}\int_0^{\pi}\cdots\int_0^{\pi} 1_E(v(r,\phi_1,\ldots,\phi_{k-1}))\sin^{k-2}\phi_1\\
\times\cdots\times\sin\phi_{k-2}\,d\phi_1\cdots d\phi_{k-1}
\end{aligned}
$$

for $E=C$ or $E=H$ and $v(r,\phi_1,\ldots,\phi_{k-1})$ the point in $\mathbb{R}^k$ having polar coordinates $r,\phi_1,\ldots,\phi_{k-1}$. The second equation is derived by the chain of identities

$$
\begin{aligned}
\frac{2^{(1-k)/2}}{\Gamma(k/2)}\int_0^\infty r^{k-1}\,\mu_{k,r}(H)\,e^{-r^2/2}\,dr&=N(H)=\Phi(s)=F(t)\\
&=N\{x \mid \|x\|\leqslant t\}\\
&=\frac{2^{(1-k)/2}}{\Gamma(k/2)}\int_0^\infty r^{k-1}\,\mu_{k,r}(C)\,e^{-r^2/2}\,dr.
\end{aligned}
$$

As the last integrand in (16) is nonnegative, (16) is proved. From (13), (14), and (16) one derives immediately

$$tF'(t)\geqslant s\Phi'(s)>0 \tag{17}$$

and this yields

$$F^{-1}(r)(d/dr)\,\Phi^{-1}(r)=t/\Phi'(s)\geqslant s/F'(t)=\Phi^{-1}(r)(d/dr)\,F^{-1}(r). \tag{18}$$

Thus

$$(d/dr)(\Phi^{-1}/F^{-1})(r)\geqslant 0,\qquad \forall r\in(\tfrac{1}{2},1). \tag{19}$$

This concludes the proof of the theorem.

*Remark* 1. The functions $r \to \mu_{k,r}(H)$ and $r \to \mu_{k,r}(C)$ cross at most once even if $H$ does not contain the origin as an interior point (see [5]). However, the general proof is much more difficult than that one given in [4] for the special case used above. An alternative and perhaps somewhat simpler proof of the assertion in the case $0 \in H^0$ looks as follows: Suppose $\mu_{k,s}(H) < \mu_{k,s}(C)$ and let $U^{\Delta}_{s,\varepsilon}(B)$ be the closed geodesic ε-neighbourhood for a set $B \subset sS^{k-1}$. If we take $H = \{x \in \mathbb{R}^k | x_1 \leqslant b\}$, we obtain for $0 < b < s$ and $r \leqslant s$ $U^{\Delta}_{s,\alpha}(H \cap sS^{k-1}) = (s/r)H \cap sS^{k-1}$, where $\alpha = s(\arccos(b/s) - \arccos(b/r \wedge 1))$. $\alpha$ is a (strictly) increasing function of $b$. Therefore, if $\tilde{H}$ is a half space containing $C$, $U^{\Delta}_{s,\alpha}(\tilde{H} \cap sS^{k-1}) \subset (s/r)\tilde{H} \cap sS^{k-1}$. Now, if $r \leqslant s$ and $C = \bigcap_{n=1}^{\infty} H_n$ with half spaces $H_n$, we get

$$\begin{aligned}\mu_{k,r}(H) = \mu_{k,s}\left(\frac{s}{r}H\right) &= \mu_{k,s}(U^{\Delta}_{s,\alpha}(H \cap sS^{k-1}))\\ &\leqslant \mu_{k,s}\,(U^{\Delta}_{s,\alpha}(C \cap sS^{k-1}))\\ &\leqslant \mu_{k,s}\left(\bigcap_{n=1}^{\infty} U^{\Delta}_{s,\alpha}(H_n \cap sS^{k-1})\right).\end{aligned}$$

The first inequality follows from the Brunn–Minkowski theorem (see [3, p. 92], for an elegant short proof) and the second from inclusion. The last set is contained in $\bigcap_{n=1}^{\infty}((s/r)\,H_n \cap sS^{k-1}) = (s/r)\,C \cap sS^{k-1}$, and its $\mu_{k,s}$-measure thus does not exceed $\mu_{k,r}(C)$. This argument completes the proof.

*Remark* 2. The existence of a constant $c_p$ as in (1′) already follows from the fundamental inequality of [2] which states that for a normal random vector $X$ and $s > 0$ such that $P(\|X\| \leqslant s) > \frac{1}{2}$ we have

$$P(\|X\| > u) \leqslant P(\|X\| \leqslant s) \exp\left\{-\frac{u^2}{24s^2}\log\frac{P(\|X\| \leqslant s)}{P(\|X\| > s)}\right\}, \qquad u \geqslant s. \quad (20)$$

From this result an estimate of the form (1′) for some $\tilde{c}_p$ can be derived in the following way: Without restriction of generality we can assume that $P(\|X\| \leqslant 1) =: q > \frac{1}{2}$. Then it follows from (20) that $\int \|x\|\, N_A(dx) \geqslant 1 - q$ and

$$\begin{aligned}\left(\int \|x\|^p\, N_A(dx)\right)^{1/p} &\leqslant \left(1 + qp\int_1^{\infty} u^{p-1}\exp\left\{-u^2\left(\log\frac{q}{1-q}\right)\Big/24\right\}du\right)^{1/p}\\ &\leqslant \frac{1}{1-q}\left(1 + \frac{qp}{2}\Gamma\left(\frac{p}{2}\right)24^{p/2}\left(\log\frac{q}{1-q}\right)^{-p/2}\right)^{1/p}\\ &\quad\times\int \|x\|\, N_A(dx), \qquad (21)\end{aligned}$$

where $q \in (\frac{1}{2}, 1)$ is arbitrary. If we define $\tilde{c}_{p,q}$ to be the constant in (21), we obtain

$$\lim_{p\to\infty} \frac{\tilde{c}_{p,q}}{c_p} = \sqrt{\frac{6}{\pi}} \frac{1}{1-q} \left(\log \frac{q}{1-q}\right)^{-1/2} > 4.9735.$$

This shows that, for large $p, c_p$ is a much smaller constant than $\tilde{c}_{p,q}$. This is also true for small values of $p$. For example, if $p \approx 1$, $c_p \leqslant \frac{1}{5} \inf\{\tilde{c}_{p,q} | q \in (\frac{1}{2}, 1)\}$. By simple examples (e.g., (4) with $k = 1$, $\alpha_1 = p = 1$) it is seen that (1′) becomes false for $\tilde{c}_p < \frac{1}{2} c_p$. Perhaps our constant can be improved by a factor $\frac{1}{2}$.

## References

1. E. F. Beckenbach and R. Bellman, "Inequalities," Springer-Verlag, Berlin/Heidelberg/New York, 1965.
2. X. Fernique, Inegrabilité des vecteurs gaussiens, *C. R. Acad. Sci. Paris Sér. A* **270** (1970), 1698–1699.
3. T. Figiel, J. Lindenstrauss, and V. D. Milman, The dimension of almost spherical sections of convex bodies, *Acta Math.* **139** (1977), 53–94.
4. H. J. Landau and L. A. Shepp, On the supremum of a gaussian process, *Sankhyā Ser. A* **32A** (1970), 369–378.
5. W. Stadje, "Rotationssymmetrische Verteilungen und Ungleichungen für die mehrdimensionale Normalverteilung," Habilitationsschrift, Göttingen, 1980.