The Annals of Probability
1980, Vol. 8, No. 2, 262–280

MATRIX NORMALIZATION OF SUMS OF RANDOM VECTORS IN THE DOMAIN OF ATTRACTION OF THE MULTIVARIATE NORMAL[1]

BY MARJORIE G. HAHN AND MICHAEL J. KLASS

University of California, Berkeley

Let S_n be a sequence of partial sums of mean zero purely d-dimensional i.i.d. random vectors. Necessary and sufficient conditions are given for the existence of matrices A_n such that the transform of S_n by A_n is asymptotically multivariate normal with identity covariance matrix. This is more general than previous d-dimensional results. Examples are given to illustrate the need for the present approach. The matrices A_n take a particularly simple form because of a degree of uncorrelatedness between certain pairs of 1-dimensional random variables obtained by projection.

0. Introduction. For a sequence of independent and identically distributed (i.i.d.) 1-dimensional random variables Z_i, not assumed to possess finite second moments, Lévy and Feller obtained necessary and sufficient conditions for the existence of constants a_n such that $\Sigma_{i=1}^n Z_i / a_n$ converges weakly to the standard normal random variable. We are interested in the central limiting behavior of $\mathbb{R}^d$-valued random vectors, especially those without finite second moments. This problem has received little attention since it seemed to require only straightforward extension of the 1-dimensional case via application of the Cramèr–Wold device. Indeed, for vectors Z_i, whose components have finite variances, scalar norming $\Sigma_{i=1}^n Z_i$ by $a_n = n^{\frac{1}{2}}$ does yield a multivariate normal limit. In general, however, norming by scalars is not appropriate in higher dimensions because any such sequence of scalars must have the same order of magnitude as the maximum of all the componentwise 1-dimensional norming constants. The limit will be degenerate for all components whose normalizing constants are of lower order than the maximum. This results in considerable loss of information and thereby prevents accurate approximation of joint distributions in $\mathbb{R}^d$

As an alternative, it is reasonable to consider componentwise normalization, in effect using d-dimensional norming constants. However, this type of norming fails for random vectors whose components have truncated pairwise correlations which do not converge (see Examples 2 and 4), or which converge to a singular limiting covariance matrix (see Example 1).

Received August 17, 1977; revised September 11, 1978.

[1]This research was supported by the National Science Foundation under grant number MCS-75-10376-AO1.

AMS 1970 *subject classifications.* Primary 60F05.

Key words and phrases. Central limit theorem, truncated correlation, infinite variance, matrix normalization, multivariate normal, random vectors.

To remedy these difficulties, we use matrices for normalization. Specifically, we find necessary and sufficient conditions for the existence of $d \times d$ matrices A_n such that the random vectors obtained by pre-multiplying the transpose of $\Sigma_{i=1}^{n} Z_i$ by A_n converge weakly to the standard multivariate normal. Thus we avoid the degeneracy inherent in scalar normalization and obtain convergence for a larger class of laws.

The method of proof gives a construction of the normalizing matrices A_n. They are not arbitrary but assume a rather special form. For instance, in the 2-dimensional case the matrices A_n may be selected to be the composition of a diagonal matrix with a rotation, defined as follows: first we specify a canonical method for finding 1-dimensional norming constants. Then among all lines through the origin, we choose the one for which, at the nth stage, the norming constant for the 1-dimensional random variable obtained by projection on this line is the smallest. This line determines what we call the "nth minimal direction." Next we truncate the random variable in the "nth minimal direction" at its nth norming constant and use the remarkable fact that it is uncorrelated with the random variable obtained by projection on the perpendicular subspace. After rotating the plane to bring these two directions to the x- and y-axes, componentwise norming does the job. The necessary and sufficient condition for matrix norming to work is merely an application of the 1-dimensional criterion uniformly in all directions.

In Section 1, we state a slight generalization of the classical normal convergence criterion and establish some asymptotic relationships needed in the rest of the paper. Section 2 treats the 2-dimensional case as described above. The generalization of $\mathbb{R}^d$ follows in Section 3. In Section 4 we investigate the relationship between norming constants in various directions. Examples are provided in Section 5, the first two of which illustrate the failure of various norming methods. Example 4 is of particular interest. It shows that the uniform sufficient condition is genuinely needed and may fail to hold even when the 1-dimensional condition is satisfied in each direction separately. This situation is new. It does not occur when each component of the random vector has finite variance.

1. Preliminaries. We assume familiarity with the classical normal convergence criterion for triangular arrays. The following formulation will be the most useful for our purposes (see e.g., Loève, Sec. 22.5).

Normal convergence criterion. If $\{X_{n,k}, k = 1, \cdots, j_n\}$ are i.i.d. random variables, then

$$\mathcal{L}\left(\Sigma_{k=1}^{j_n} X_{n,k}\right) \to N(0, 1)$$

if, and only if, for every $\varepsilon > 0$ and a $\tau > 0$,

(i) $\lim_{n\to\infty} j_n P(|X_{n,1}| > \varepsilon) = 0$;

(ii) $\lim_{n\to\infty} j_n E X_{n,1}^2 I_{(|X_{n,1}| \leqslant \tau)} = 1$;

(iii) $\lim_{n\to\infty} j_n E X_{n,1} I_{(|X_{n,1}| \leqslant \tau)} = 0$.

This formulation does not include the u.a.n. condition, $\lim_{n\to\infty} P(|X_{n,1}| > \varepsilon) = 0$ for each $\varepsilon > 0$, in the "only if" direction because it is automatically satisfied for row-wise i.i.d. random variables whose sums converge in law.

We need the following slight generalization of this condition.

PROPOSITION 1. *Let J be a nonempty set. Suppose that for each $n \geq 1$ and $\alpha \in J$, $X_{n,1,\alpha}, \cdots, X_{n,j_n,\alpha}$ are* i.i.d. *random variables. Then for every sequence $\{\alpha_n, n = 1, 2, \cdots\}$ from J*

$$\mathcal{L}\left(\Sigma_{k=1}^{j_n} X_{n,k,\alpha_n}\right) \to N(0, 1)$$

if, and only if, for every $\varepsilon > 0$ and a $\tau > 0$,

(i) $\lim_{n\to\infty} \sup_{\alpha\in J} j_n P(|X_{n,1,\alpha}| > \varepsilon) = 0$;

(ii) $\lim_{n\to\infty} \sup_{\alpha\in J} |j_n EX_{n,1,\alpha}^2 I_{(|X_{n,1,\alpha}|\leq\tau)} - 1| = 0$;

(iii) $\lim_{n\to\infty} \sup_{\alpha\in J} j_n |EX_{n,1,\alpha} I_{(|X_{n,1,\alpha}|\leq\tau)}| = 0$.

Observe that (i) and (ii) are equivalent to (i) and

(ii)′ $\lim_{n\to\infty} \sup_{\alpha\in J} |j_n E(X_{n,1,\alpha}^2 \wedge \tau^2) - 1| = 0$.

In order to simplify notation, for a one-parameter family of random variables $\{Y_\alpha, \alpha \in J\}$ let

$$L_{Y_\alpha}(t) \equiv L_\alpha(t) = EY_\alpha^2 I_{(|Y_\alpha|<t)}$$

and

$$M_{Y_\alpha}(t) \equiv M_\alpha(t) = E\left(Y_\alpha^2 \wedge t^2\right).$$

Whenever J is a singleton, the subscript α on L and M will be suppressed.

Let $Y_{\alpha,k}$, $k = 1, 2, \cdots$ be i.i.d. and $a_n(\alpha)$ a sequence of positive constants. If $X_{n,k,\alpha} = Y_{\alpha,k}/a_n(\alpha)$ then the following lemma will be helpful in checking (i)–(iii) of Proposition 1.

EQUIVALENCE LEMMA. *For a one-parameter family of random variables $\{Y_\alpha, \alpha \in J\}$ the following are equivalent*:

(1) $\lim_{t\to\infty} \sup_{\alpha\in J} t^2 P(|Y_\alpha| > t)/M_\alpha(t) = 0$;

(2) *for every $\varepsilon > 0$ and all $0 \leq \beta < 2$,*

$$\lim_{t\to\infty} \sup_{\alpha\in J} t^{2-\beta} E|Y_\alpha|^\beta I_{(|Y_\alpha|>\varepsilon t)}/M_\alpha(t) = 0;$$

(3) $\lim_{t\to\infty} \sup_{\alpha\in J} M_\alpha(2t)/M_\alpha(t) = 1$;

(4) $\lim_{t\to\infty} \sup_{\alpha\in J} M_\alpha(t)/L_\alpha(t) = 1$.

PROOF. Assuming (1),

$$\begin{aligned} 0 &\leq M_\alpha(2t)/M_\alpha(t) - 1 \\ &= 2\textstyle\int_t^{2t} uP(|Y_\alpha| > u)du/M_\alpha(t) \\ &\leq 2P(|Y_\alpha| > t)\textstyle\int_t^{2t} udu/M_\alpha(t) \\ &= 3t^2 P(|Y_\alpha| > t)/M_\alpha(t) \to 0 \text{ uniformly in } \alpha. \end{aligned}$$

Hence, (3) is true.

Now, assuming (3),

$$\begin{aligned}\lim_{t\to\infty} \sup_{\alpha\in J} t^2P(|Y_\alpha| > t)/M_\alpha(t) &\leqslant \lim_{t\to\infty} \sup_{\alpha\in J} 8\textstyle\int_{t/2}^{t} uP(|Y_\alpha| > u)du/3M_\alpha(t)\\ &\leqslant \lim_{t\to\infty} \sup_{\alpha\in J} 4(M_\alpha(t) - M_\alpha(t/2))/3M_\alpha(t)\\ &= 0.\end{aligned}$$

Consequently, (1) and (3) are equivalent.

Clearly (2) $\Rightarrow$ (1). Assuming both (1) and (3) we will prove (2). Fix $0 \leqslant \beta < 2$. Let $0 < \varepsilon \leqslant 1$, $C > 4/(2^{2-\beta} - 1)$, $0 < \delta < 2^{2-\beta} - 1 - 4/C$, and $\gamma = 1 + \delta$. There exists T such that $t \geqslant T$ implies

$$\sup_{\alpha\in J} M_\alpha(2t)/M_\alpha(t) \leqslant \gamma \tag{1.1}$$

and

$$\sup_{\alpha\in J} t^2P(|Y_\alpha| > \varepsilon t)/M_\alpha(t) \leqslant \delta. \tag{1.2}$$

Note that (1.2) follows since $M_\alpha(\varepsilon t) \leqslant M_\alpha(t)$. Thus, assume $\beta > 0$. Then, for any $\alpha \in J$ and $t \geqslant T$,

$$\begin{aligned}t^{2-\beta}E|Y_\alpha|^\beta I_{(|Y_\alpha|>\varepsilon t)} &= -t^{2-\beta}u^\beta P(|Y_\alpha| > u)|_{u=\varepsilon t}^{\infty} + t^{2-\beta}\beta\textstyle\int_{\varepsilon t}^{\infty} u^{\beta-1}P(|Y_\alpha| > u)du\\ &\leqslant t^2\varepsilon^\beta P(|Y_\alpha| > \varepsilon t) + t^{2-\beta}\beta\textstyle\sum_{n=1}^{\infty}\int_{2^{n-1}\varepsilon t}^{2^n\varepsilon t} u^{\beta-1}P(|Y_\alpha| > u)du\\ &\leqslant \varepsilon^\beta\delta M_\alpha(t) + t^{2-\beta}\textstyle\sum_{n=1}^{\infty}(2^n\varepsilon t)^\beta P(|Y_\alpha| > 2^{n-1}\varepsilon t) \qquad \text{by (1.2)}\\ &\leqslant \varepsilon^\beta\delta M_\alpha(t) + \delta\textstyle\sum_{n=1}^{\infty}(2\varepsilon)^\beta(2^{n-1})^{\beta-2}M_\alpha(2^{n-1}t) \qquad \text{by (1.2)}\\ &\leqslant \varepsilon^\beta\delta M_\alpha(t)\big(1 + 2^\beta\textstyle\sum_{n=1}^{\infty}(\gamma/2^{2-\beta})^{n-1}\big) \qquad \text{by (1.1)}\\ &= \varepsilon^\beta\delta M_\alpha(t)(1 + 2^\beta/(1 - \gamma/2^{2-\beta}))\\ &\leqslant (C + 1)\varepsilon^\beta\delta M_\alpha(t).\end{aligned}$$

Hence,

$$\lim_{t\to\infty} \sup_{\alpha\in J} t^{2-\beta}E|Y_\alpha|^\beta I_{(|Y_\alpha|>\varepsilon t)}/M_\alpha(t) \leqslant (C + 1)\varepsilon^\beta\delta.$$

Since $\delta > 0$ is arbitrary, (2) holds.

The equivalence of (4) and (1) follows immediately from

$$L_\alpha(t)/M_\alpha(t) = 1 - t^2P(|Y_\alpha| > t)/M_\alpha(t).$$

Hence the equivalence of (1)-(4). □

Feller ((1966), pages 312-13) was the first to notice that a random variable belongs to the domain of attraction of the normal if and only if $L(t)$ is a slowly varying function, i.e., for each $s > 0$, $\lim_{t\to\infty} L(st)/L(t) = 1$. In order to verify that an increasing function g is slowly varying it suffices to show that

$$\lim_{t\to\infty} g(2t)/g(t) = 1,$$

because if $1 \leqslant 2^{n-1} \leqslant s \leqslant 2^n$, for example, then

$$1 = \lim_{t\to\infty} \Pi_{k=1}^{n-1} g(2^k t)/g(2^{k-1}t) \leqslant \lim_{t\to\infty} g(st)/g(t) \leqslant \lim_{t\to\infty} \Pi_{k=1}^{n} g(2^k t)/g(2^{k-1}t) = 1.$$

Therefore, by the equivalence lemma, X is in the domain of attraction of the normal if and only if M is slowly varying, a condition which is in turn equivalent to

$$\lim_{y\to\infty} y^2 P(|X| > y)/M(y) = 0. \tag{1.3}$$

As a final remark, notice that if the triangular array $X_{n,k} = Y_k/a_n$, $k = 1, \cdots, n$, satisfies the normal convergence criterion with $Y_1, Y_2, \cdots$ i.i.d. random variables, then the norming constants a_n are determined up to asymptotic equivalence by

$$\lim_{n\to\infty} nL(a_n)/a_n^2 = 1.$$

By the equivalence lemma, the a_n's also satisfy

$$\lim_{n\to\infty} nM(a_n)/a_n^2 = 1.$$

2. Central limit theorem in $\mathbb{R}^2$. Given a random vector $Z = (X, Y)$ in $\mathbb{R}^2$, the perpendicular projection of Z on the line of angle θ through the origin is defined by

$$X_\theta = X\cos\theta + Y\sin\theta.$$

The random vector Z is purely 2-dimensional if and only if $P(X_\theta \neq 0) > 0$ for all $\theta \in [0, 2\pi]$. This is equivalent to $0 < EX_\theta^2 \leqslant \infty$ for all θ. If W is either a column or row vector, let ${}^tW$ denote its transpose.

The main result of this paper is

THEOREM 2. *Let $(X_1, Y_1), (X_2, Y_2), \cdots$ be* i.i.d. *purely 2-dimensional mean zero random vectors. There exist matrices A_n such that*

$${}^t\left(A_n \Sigma_{i=1}^n {}^t(X_i, Y_i)\right) \to_{\mathcal{D}} N(\vec{0}, I),$$

where I is the 2×2 identity matrix, if and only if,

$$\lim_{t\to\infty} \sup_\theta t^2 P(|X_\theta| > t)/M_\theta(t) = 0. \tag{2.1}$$

Moreover, let $a_n(\theta)$ be the largest real number satisfying $a_n^2(\theta) = nM_\theta(a_n(\theta))$. There exists $0 \leqslant \gamma_n < \pi$ such that $a_n(\gamma_n) = \inf_\theta a_n(\theta)$. When (2.1) holds we may let

$$A_n = \begin{pmatrix} (\cos\gamma_n)/a_n(\gamma_n) & (\sin\gamma_n)/a_n(\gamma_n) \\ -(\sin\gamma_n)/a_n(\gamma_n + \pi/2) & (\cos\gamma_n)/a_n(\gamma_n + \pi/2) \end{pmatrix}.$$

PROOF OF NECESSITY. Let the norming matrices A_n be denoted by

$$A_n = \begin{pmatrix} r_n & t_n \\ u_n & v_n \end{pmatrix}.$$

Let $c_n^2 = (r_n^2 + t_n^2)^{-1}$and $\tilde{c}_n^2 = (u_n^2 + v_n^2)^{-1}$. Let θ_n and $\tilde{\theta}_n$ satisfy

$$\cos\theta_n = r_n c_n, \sin\theta_n = t_n c_n$$
$$\cos\tilde{\theta}_n = u_n \tilde{c}_n, \sin\tilde{\theta}_n = v_n \tilde{c}_n.$$

Assuming normal convergence,

$$\begin{aligned} &{}^t\big(\Sigma_{i=1}^n (X_{\theta_n})_i / c_n, \Sigma_{i=1}^n (X_{\tilde{\theta}_n})_i / \tilde{c}_n\big) \\ &= {}^t\big(\Sigma_{i=1}^n (X_i \cos\theta_n + Y_i \sin\theta_n)/c_n, \Sigma_{i=1}^n (X_i \cos\tilde{\theta}_n + Y_i \sin\tilde{\theta}_n)/\tilde{c}_n\big) \\ &= {}^t(r_n \Sigma_{i=1}^n X_i + t_n \Sigma_{i=1}^n Y_i, u_n \Sigma_{i=1}^n X_i + v_n \Sigma_{i=1}^n Y_i) \\ &= {}^t\big(A_n \Sigma_{i=1}^n {}^t(X_i, Y_i)\big) \to N(\vec{0}, I). \end{aligned}$$

At this point, in order to obtain (2.1), we require a strengthened version of usual weak convergence results. The following lemma, due to R. Rao, can be found in Billingsley ((1968), Problem 8, Page 17).

LEMMA. *Let S be a separable metric space and $\{f_\theta\}$ a uniformly bounded family of real-valued functions defined on S, equicontinuous at each x (for each x and $\varepsilon > 0$ there exists $\delta > 0$ for which $\rho(x, y) < \delta$ implies $\sup_\theta |f_\theta(x) - f_\theta(y)| < \varepsilon$). Let P_n and P be probability measures on S such that P_n converges weakly to P. Then*

$$\lim_{n\to\infty} \sup_\theta \left|\int f_\theta dP_n - \int f_\theta dP\right| = 0. \tag{2.2}$$

Since this result holds for pairs $\{g_\theta\}$ and $\{h_\theta\}$ of uniformly bounded real-valued equicontinuous families of functions on S, it holds if f_θ is complex-valued. We apply the lemma with $S = \mathbb{R}^2$,

$$f_\theta(\vec{x}) = f_\theta(x_1, x_2) = \exp\{iu(x_1\cos\theta + x_2\sin\theta)\},$$
$$P_n = \mathcal{L}\big(\Sigma_{i=1}^n (X_{\theta_n})_i / c_n, \Sigma_{i=1}^n (X_{\tilde{\theta}_n})_i / \tilde{c}_n\big),$$

and

$$P = N(\vec{0}, I).$$

The family $\{f_\theta, 0 \leqslant \theta \leqslant 2\pi\}$ is equicontinuous at $\vec{x}$. Hence, for each $u \in \mathbb{R}$,

$$\lim_{n\to\infty} \sup_{0\leqslant\theta\leqslant 2\pi} \left| E\exp\left\{ iu\left(\frac{\cos\theta}{c_n} \Sigma_{j=1}^n (X_{\theta_n})_j + \frac{\sin\theta}{\tilde{c}_n} \Sigma_{j=1}^n (X_{\tilde{\theta}_n})_j \right) \right\} - e^{-u^2/2} \right| = 0.$$

By the Lévy continuity theorem, for each sequence of reals $\{\psi_n\}$ with $0 \leqslant \psi_n \leqslant 2\pi$,

$$\mathcal{L}\left(\frac{\cos\psi_n}{c_n} \Sigma_{j=1}^n (X_{\theta_n})_j + \frac{\sin\psi_n}{\tilde{c}_n} \Sigma_{j=1}^n (X_{\tilde{\theta}_n})_j \right) \to N(0, 1).$$

Since (for all n sufficiently large), X_{θ_n} and $X_{\tilde{\theta}_n}$ are not linearly dependent, for each $0 \leqslant \phi \leqslant 2\pi$, there exist unique constants $\beta_n(\phi)$ and $\tilde{\beta}_n(\phi)$ such that

$$X_\phi = \beta_n(\phi) X_{\theta_n} / c_n + \tilde{\beta}_n(\phi) X_{\tilde{\theta}_n} / \tilde{c}_n.$$

Let $b_n^2(\phi) = \beta_n^2(\phi) + \tilde{\beta}_n^2(\phi)$. This is positive, so there exists $0 \leqslant g_n(\phi) \leqslant 2\pi$ such that $\cos g_n(\phi) = \beta_n(\phi)/b_n(\phi)$ and $\sin g_n(\phi) = \tilde{\beta}_n(\phi)/b_n(\phi)$. Hence, for each sequence $\{\phi_n\}$,

$$\Sigma_{j=1}^n (X_{\phi_n})_j/b_n(\phi_n) = \cos g_n(\phi_n)\Sigma_{j=1}^n(X_{\theta_n})_j/c_n + \sin g_n(\phi_n)\Sigma_{j=1}^n(X_{\tilde{\theta}_n})_j/\tilde{c}_n \to_{\mathfrak{D}} N(0, 1).$$

Applying Proposition 1 with $X_{n,k,\phi} = (X_\phi)_k/b_n(\phi)$ gives

$$\lim_{n\to\infty} \sup_\phi nP(|X_\phi| > \varepsilon b_n(\phi)) = 0 \quad \text{for each} \quad \varepsilon > 0, \tag{2.3}$$

and

$$\lim_{n\to\infty} \sup_\phi |nM_\phi(b_n(\phi))/b_n^2(\phi) - 1| = 0. \tag{2.4}$$

We assert the validity of (2.1). If not, there exists $\delta > 0$, $0 \leqslant \phi_n \leqslant 2\pi$ and $t_n \to \infty$ such that

$$t_n^2 P(|X_{\phi_n}| > t_n)/M_{\phi_n}(t_n) \geqslant 3\delta.$$

Take integers j_n so that

$$\lim_{n\to\infty} 2j_n M_{\phi_n}(t_n)/t_n^2 = 1. \tag{2.5}$$

Then $j_n P(|X_{\phi_n}| > t_n) \geqslant \delta$ for n sufficiently large. However, since $M_\phi(t)/t^2 = E((X_\phi/t)^2 \wedge 1)$ decreases as t increases, (2.4) and (2.5) imply that $b_{j_n}(\phi_n) < t_n$ for n sufficiently large. Referring to (2.3),

$$j_n P(|X_{\phi_n}| > t_n) \leqslant j_n P(|X_{\phi_n}| > b_{j_n}(\phi_n)) \to 0,$$

which gives the desired contradiction. Hence, (2.1) holds. □

In order to prove sufficiency it is convenient to have a canonical definition of the norming constants $a_n(\theta)$. Let $y_\theta = \sup\{y \geqslant 0: P(|X_\theta| \geqslant y) = P(|X_\theta| > 0)\}$. Clearly, since $P(|X_\theta| > 0) > 0$, $M_\theta(y)/y^2$ is strictly decreasing and continuous on $[y_\theta, \infty)$, varying from $P(X_\theta \neq 0)$ to zero. Therefore, for each $u \geqslant 1/P(X_\theta \neq 0)$ there exists a unique $a(u, \theta)$ with $y_\theta \leqslant a(u, \theta) < \infty$ such that

$$\frac{1}{u} = M_\theta(a(u, \theta))/a^2(u, \theta).$$

It is easily shown that $u^* \equiv \sup_\theta 1/P(X_\theta \neq 0) < \infty$. Hence, as the canonical norming constants we choose $a_n^2(\theta) \equiv a^2(n, \theta)$ for $n \geqslant u^*$. In the sequel we always assume $n > u^*$.

One nice property of this definition is that $\theta \to a_n(\theta)$ is continuous: note that $X_\theta \to X_{\theta_0}$ as $\theta \to \theta_0$. Hence, an application of the dominated convergence theorem shows that $M_\theta(y)y^{-2}$ is, in fact, jointly continuous in y and θ. Due to this and the strict monotonicity of $M_\theta(y)y^{-2}$ for $y \geqslant y_\theta$, for each fixed $\theta^* n > u^*$, and sufficiently small $\delta \equiv \delta(\theta^*, n) > 0$; there exists $\varepsilon > 0$ such that $|\theta - \theta^*| < \delta$ and $|y - a_n(\theta^*)| > \delta$ together imply

$$G_{n,\theta}(y) \equiv \left|M_\theta(y)y^{-2} - \frac{1}{n}\right| = |M_\theta(y)y^{-2} - M_\theta^*(a_n(\theta^*))/a_n^2(\theta^*)| > \varepsilon.$$

Since $G_{n,\theta}(a_n(\theta)) = 0$, it follows that $\theta \to \theta^*$ implies $a_n(\theta) \to a_n(\theta^*)$, proving the asserted continuity.

PROOF OF SUFFICIENCY. Let $a_n(\theta)$ be the canonically defined norming constants for X_θ. Note that there exists $\gamma_n \in [0, \pi]$ such that $C_n \equiv \inf_\theta a_n(\theta) = a_n(\gamma_n) > 0$. This follows from the compactness of $[0, 2\pi]$, the continuity of $\theta \to a_n(\theta)$, and the fact that $a_n(\theta) > 0$ for $n > u^*$. Since $a_n(\theta) \geqslant (n/n^*)^{\frac{1}{2}} a_n^*(\theta) \geqslant C_n(n/n^*)^{\frac{1}{2}}$, we conclude that $\lim_{n\to\infty} \inf_\theta a_n(\theta) = \infty$. In view of the equivalance lemma, (2.1.) now implies

$$\lim_{n\to\infty} \sup_\theta nP(|X_\theta| > \varepsilon a_n(\theta)) = 0 \qquad \text{for all} \quad \varepsilon > 0 \tag{2.6}$$

and in fact the stronger statement

$$\lim_{n\to\infty} \sup_\theta nE|X_\theta| I_{(|X_\theta| > \varepsilon a_n(\theta))}/a_n(\theta) = 0 \qquad \text{for all} \quad \varepsilon > 0. \tag{2.7}$$

Since X_{γ_n} uses the minimal canonical norming constant at stage n, so we call γ_n the "nth minimal direction". Let $\tilde{\gamma}_n = \gamma_n + \pi/2$.

We suppress the dependence on n and write

$$X_\theta = X_\theta I_{(|X_\theta| < a_n(\theta))} + X_\theta I_{(|X_\theta| > a_n(\theta))}$$
$$\equiv X_\theta' + X_\theta''.$$

Let

$$U_n = \Sigma_{j=1}^n (X_{\gamma_n})_j, \qquad U_n' = \Sigma_{j=1}^n (X_{\gamma_n}')_j,$$
$$V_n = \Sigma_{j=1}^n (X_{\tilde{\gamma}_n})_j, \qquad V_n' = \Sigma_{j=1}^n (X_{\tilde{\gamma}_n}')_j.$$

The proof of sufficiency will be completed once it is shown that

$$\mathcal{L}(U_n'/a_n(\gamma_n), V_n'/a_n(\tilde{\gamma}_n)) \to N(\vec{0}, I). \tag{2.8}$$

To see that this suffices, define

$$A_n = \begin{pmatrix} 1/a_n(\gamma_n) & 0 \\ 0 & 1/a_n(\tilde{\gamma}_n) \end{pmatrix} \begin{pmatrix} \cos\gamma_n & \sin\gamma_n \\ -\sin\gamma_n & \cos\gamma_n \end{pmatrix}.$$

Clearly,

$$A_n \Sigma_{j=1}^n {}^t(X_j, Y_j) = {}^t(U_n/a_n(\gamma_n), V_n/a_n(\tilde{\gamma}_n))$$

and by (2.6),

$$P((U_n/a_n(\gamma_n), V_n/a_n(\tilde{\gamma}_n)) \neq (U_n'/a_n(\gamma_n), V_n'/a_n(\tilde{\gamma}_n)))$$
$$\leqslant nP(|X_{\gamma_n}| > a_n(\gamma_n)) + nP(|X_{\tilde{\gamma}_n}| > a_n(\tilde{\gamma}_n)) \to 0 \qquad \text{as} \quad n \to \infty.$$

Therefore,

$$\lim_{n\to\infty} \mathcal{L}(U_n/a_n(\gamma_n), V_n/a_n(\tilde{\gamma}_n)) = \lim_{n\to\infty} \mathcal{L}(U_n'/a_n(\gamma_n), V_n'/a_n(\tilde{\gamma}_n)).$$

Weak convergence in $\mathbb{R}^k$ holds if and only if all fixed linear combinations of components converge appropriately (see Billingsley (1968), Theorem 7.7, Page 49). Scaling, in order to prove (2.8), it suffices to show that for all fixed θ,

$$U_n'(\cos\theta)/a_n(\gamma_n) + V_n'(\sin\theta)/a_n(\tilde{\gamma}_n) \to_{\mathfrak{D}} N(0, 1). \tag{2.9}$$

Let $W_n(\theta) = X_{\gamma_n}'(\cos\theta)/a_n(\gamma_n) + X_{\tilde{\gamma}_n}'(\sin\theta)/a_n(\tilde{\gamma}_n)$. Note that $I_{(|W_n(\theta)| \leqslant 2)} = 1$ a.s.

Hence, invoking the normal convergence criterion, we require the three conditions

(2.10) $\lim_{n\to\infty} nP(|W_n(\theta)| > \varepsilon) = 0 \quad$ for every $\varepsilon > 0$

(2.11) $\lim_{n\to\infty} nEW_n^2(\theta) = 1$

(2.12) $\lim_{n\to\infty} nEW_n(\theta) = 0.$

Now,

$$nP(|W_n(\theta)| > \varepsilon) \leqslant nP\big(|X_{\gamma_n}| > \varepsilon a_n(\gamma_n)/2\big) + nP\big(|X_{\tilde{\gamma}_n}| > \varepsilon a_n(\tilde{\gamma}_n)/2\big) \to 0$$

by (2.6), which proves (2.10).

Since $EX_\theta = 0$, $EX'_\theta = -EX''_\theta$. Therefore,

$$n|EW_n(\theta)| \leqslant nE|X_{\gamma_n}|I_{(|X_{\gamma_n}|>a_n(\gamma_n))}/a_n(\gamma_n) + nE|X_{\tilde{\gamma}_n}|I_{(|X_{\tilde{\gamma}_n}|>a_n(\tilde{\gamma}_n))}/a_n(\tilde{\gamma}_n) \to 0$$

by (2.7). This verifies (2.12).

Finally,

$$\begin{aligned} nEW_n^2(\theta) &= n\cos^2\theta\big(M_{\gamma_n}(a_n(\gamma_n)) - a_n^2(\gamma_n)P(|X_{\gamma_n}| > a_n(\gamma_n))\big)/a_n^2(\gamma_n) \\ &\quad + n\sin^2\theta\big(M_{\tilde{\gamma}_n}(a_n(\tilde{\gamma}_n)) - a_n^2(\tilde{\gamma}_n)P(|X_{\tilde{\gamma}_n}| > a_n(\tilde{\gamma}_n))\big)/a_n^2(\tilde{\gamma}_n) \\ &\quad + n\sin 2\theta EX'_{\gamma_n}X'_{\tilde{\gamma}_n}/a_n(\gamma_n)a_n(\tilde{\gamma}_n) \\ &= \cos^2\theta + \sin^2\theta + o(1) + n\sin 2\theta EX'_{\gamma_n}X'_{\tilde{\gamma}_n}/a_n(\gamma_n)a_n(\tilde{\gamma}_n) \\ &\to 1 + \sin 2\theta \lim_{n\to\infty} nEX'_{\gamma_n}X'_{\tilde{\gamma}_n}/a_n(\gamma_n)a_n(\tilde{\gamma}_n). \end{aligned}$$

To treat the above limit, we use the fact that $EX'_{\gamma_n}X_{\tilde{\gamma}_n} = 0$ which will be proved in Corollary 4 below. Hence,

$$\begin{aligned} n|EX'_{\gamma_n}X'_{\tilde{\gamma}_n}|/a_n(\gamma_n)a_n(\tilde{\gamma}_n) &= n|EX'_{\gamma_n}X_{\tilde{\gamma}_n}I_{(|X_{\tilde{\gamma}_n}|>a_n(\tilde{\gamma}_n))}|/a_n(\gamma_n)a_n(\tilde{\gamma}_n) \\ &\leqslant nE|X_{\tilde{\gamma}_n}|I_{(|X_{\tilde{\gamma}_n}|>a_n(\tilde{\gamma}_n))}/a_n(\tilde{\gamma}_n) \to 0 \end{aligned}$$

by(2.7). Thus, (2.11) holds.

This substantiates (2.8), completing the proof, modulo Corollary 4. □

Observe that condition (2.1) is simply the uniform version of condition (1.3), which was noted to be necessary and sufficient for the CLT to hold in ℝ.

Before attending to Corollary 4, we remark that the above proof only requires that the CLT hold for two sequences of random variables X_{θ_n} and X_{ξ_n} which are "asymptotically uncorrelated" in the sense that

(2.13) $\lim_{n\to\infty} nEX'_{\theta_n}X'_{\xi_n}/a_n(\theta_n)a_n(\xi_n) = 0.$

Choice of the "nth minimal direction", γ_n, for θ_n becomes natural since for the perpendicular direction, $\tilde{\gamma}_n$, $EX'_{\gamma_n}X_{\tilde{\gamma}_n} = 0$, which, in turn, easily yields "asymptotic uncorrelation". It is also possible to construct two "asymptotically uncorrelated" random variables by considering the "nth maximal direction" and its perpendicular. This is a conceptually different but, nevertheless, essentially equivalent approach.

The next lemma, on which Corollary 4 is based, is basic to our construction of "asymptotically uncorrelated" random variables. It was motivated by the observation that for random variables X and Y with finite second moments, $EX^2 = \inf_\theta EX_\theta^2$ if and only if $EXY = 0$ and $EX^2 \leqslant EY^2$.

LEMMA 3. *Let X and Y be arbitrary random variables with finite expectations. For $t \geqslant 0$, let $0 \leqslant \theta_t < \pi$ satisfy*

$$E\left(X_{\theta_t}^2 \wedge t^2\right) = \inf_\theta E\left(X_\theta^2 \wedge t^2\right).$$

Then

$$EX_{\theta_t}X_{\theta_t+\pi/2}I_{(|X_{\theta_t}|<t)} = 0. \tag{2.14}$$

PROOF. For $t = 0$ the result is trivial. Fix $t > 0$ and let $g(\theta) = E(X_\theta^2 \wedge t^2)$. We first show that such a θ_t exists. Note that $x^2 \wedge t^2$ is a bounded continuous function and $X_\theta \to X_{\theta^*}$ a.s. as $\theta \to \theta^*$. Hence $g(\theta)$ is continuous in θ. Since $[0, 2\pi]$ is compact, there exists $\theta_t \in [0, 2\pi]$ such that $g(\theta_t) = \inf_\theta g(\theta)$. Now because $g(\theta) = g(\theta + k\pi)$ for $k = \pm 1, \pm 2, \cdots$, we may take $0 \leqslant \theta_t < \pi$. Next note that

$$\{X_\theta : 0 \leqslant \theta \leqslant 2\pi\} = \left\{X_{\theta_t}\cos\phi + X_{\theta_t+\pi/2}\sin\phi : -\pi \leqslant \phi \leqslant \pi\right\}.$$

Hence, it suffices to assume $\theta_t = 0$, and operate on $-\pi \leqslant \theta \leqslant \pi$. Thus $g(0) = \inf_\theta g(\theta)$. We will show that g has both a right and a left hand derivative at $\theta = 0$.

Since $E|Y| < \infty$, $\lim_{\theta\to 0} P(|Y| > |1/\theta|)/\theta = 0$. Hence, there exists a symmetric continuous function $f(\theta)$, monotone increasing on $[0, \pi]$ with $f(0) = 0$, $f(\pi) \leqslant 1$ such that $\lim_{\theta\to 0} P(|Y\sin\theta| > f(\theta))/\theta = 0$. Let $A_\theta = \{|Y\sin\theta| \leqslant f(\theta)\}$.

We intend to consider the difference quotient $(g(\theta) - g(0))/\theta$ and apply dominated convergence. Let

$$Q(\theta) = \left(X_\theta^2 \wedge t^2 - X^2 \wedge t^2\right)/\theta.$$

Observe that attention may be restricted to the event A_θ since

$$E|Q(\theta)|I_{A_\theta^c} \leqslant t^2 P(A_\theta^c)/|\theta| \to 0 \quad \text{as} \quad \theta \to 0.$$

First of all,

(2.15)

$$\lim_{\theta\to 0} Q(\theta)I_{A_\theta}I_{(|X|<t)} = \lim_{\theta\to 0}\left(\frac{XY\sin 2\theta}{\theta} + Y(Y\sin\theta)\frac{\sin\theta}{\theta}\right)I_{A_\theta}I_{(|X|<t)}$$
$$= 2XYI_{(|X|<t)} \quad \text{a.s.}$$

Second,

$$\lim_{\theta\to 0} Q(\theta)I_{A_\theta}I_{(|X|>t)} = 0 \text{ a.s.} \tag{2.16}$$

In order to deal with pointwise convergence when $|X| = t$, we must distinguish between two cases:

CASE 1. $\theta > 0$.

(2.17) $\lim_{\theta\downarrow 0^+} Q(\theta) I_{A_\theta} I_{(|X|=t)}$

$$= \lim_{\theta\downarrow 0^+} \frac{1}{\theta}(t^2\cos^2\theta + XY\sin 2\theta + Y^2\sin^2\theta - t^2) I_{A_\theta} I_{(|X|=t)} I_{(|X_\theta|<t)}$$

$$= 2XYI_{(|X|=t,\, XY<0)} \qquad \text{a.s.}$$

CASE 2. $\theta < 0$. Similarly,

$$\lim_{\theta\uparrow 0^-} Q(\theta) I_{A_\theta} I_{(|X|=t)} = 2XYI_{(|X|=t,\, XY>0)} \qquad \text{a.s.} \tag{2.18}$$

To bound the variables, note that $|X_\theta^2 \wedge t^2 - X^2 \wedge t^2| \leqslant |X_\theta^2 - X^2|$ so that

$$|Q(\theta)| I_{A_\theta} = |Q(\theta)| I_{A_\theta} I_{(|X\cos\theta|<t+f(\theta))}$$

$$\leqslant \left(X^2\frac{\sin^2\theta}{|\theta|} + |XY\frac{\sin 2\theta}{\theta}| + |Y^2\frac{\sin^2\theta}{\theta}|\right) I_{A_\theta} I_{(|X\cos\theta|<t+f(\theta))}$$

$$\leqslant (t+1)^2 + [2(t+1)+1]|Y| \qquad \text{for } |\theta| \text{ sufficiently small,}$$

which is integrable.

Hence, using (2.15)-(2.17) and dominated convergence,

$$g'_+(0) \equiv \lim_{\theta\downarrow 0^+} (g(\theta) - g(0))/\theta = \lim_{\theta\downarrow 0^+} EQ(\theta) I_{A_\theta}$$
$$= 2EXYI_{(|X|<t)} + 2EXYI_{(|X|=t,\, XY<0)}.$$

Similarly, using (2.15), (2.16), (2.18) and dominated convergence,

$$g'_-(0) \equiv \lim_{\theta\uparrow 0^-} (g(\theta) - g(0))/\theta$$
$$= 2EXYI_{(|X|<t)} + 2EXYI_{(|X|=t,\, XY>0)}.$$

Since $g(\theta) \geqslant g(0)$ it follows that

$$g'_-(0) \leqslant 0 \leqslant g'_+(0).$$

Thus,

$$0 \leqslant g'_+(0) - g'_-(0) = -2E|XY| I_{(|X|=t)} \leqslant 0.$$

Hence, $E|XY|I_{(|X|=t)} = 0$ and $g'(0)$ exists and equals zero. Consequently, $EXYI_{(|X|\leqslant t)} = 0$, as desired. ∎

COMMENT. Letting ϕ_t satisfy $E(X_{\phi_t}^2 \wedge t^2) = \sup_\theta E(X_\theta^2 \wedge t^2)$, essentially the same proof goes through as above except that now

$$g'_+(\phi_t) \leqslant 0 \leqslant g'_-(\phi_t).$$

This implies

$$|EX_{\phi_t}X_{\phi_t+\pi/2} I_{(|X_{\phi_t}|<t)}| \leqslant tE|X_{\phi_t+\pi/2}| I_{(|X_{\phi_t}|=t)}. \tag{2.19}$$

COROLLARY 4. *Let X and Y be arbitrary random variables with finite expecta-*

tions. Let $a_n^2(\theta) = nM_\theta(a_n(\theta))$, $a_n(\gamma_n) = \inf_\theta a_n(\theta)$, *and* $\tilde{\gamma}_n = \gamma_n + \pi/2$. *Then*

$$EX'_{\gamma_n}X_{\tilde{\gamma}_n} = 0. \tag{2.20}$$

PROOF. Fix n. By continuity of $a_n(\cdot)$ and compactness of $[0, 2\pi]$, such a γ_n exists. Let $t_0 = a_n(\gamma_n)$. (2.20) holds trivially if $t_0 = 0$, so suppose $t_0 > 0$. Define $g(\theta) = M_\theta(t_0)$. Then since $a_n(\theta) \geqslant t_0 > 0$, we have

$$\begin{aligned} g(\theta) &= t_0^2 E\big((X_\theta/t_0)^2 \wedge 1\big) \\ &\geqslant t_0^2 E\big((X_\theta/a_n(\theta))^2 \wedge 1\big) \\ &= t_0^2/n \\ &= g(\gamma_n). \end{aligned}$$

Hence, $g(\gamma_n) = \inf_\theta g(\theta)$. Now apply Lemma 3. $\square$

COMMENT. When considering the "nth maximal direction", θ_n, and its perpendicular, $\xi_n = \theta_n + \pi/2$, the analogue of Corollary 4, assuming condition (2.1), becomes

$$\lim_{n\to\infty} nEX'_{\theta_n}X_{\xi_n}/a_n(\theta_n)a_n(\xi_n) = 0.$$

Using both (2.1) and (2.19),

$$\begin{aligned} &n|EX'_{\theta_n}X_{\xi_n}|/a_n(\theta_n)a_n(\xi_n) \\ &\qquad \leqslant nE|X_{\xi_n}|I_{(|X_{\theta_n}|=a_n(\theta_n))}/a_n(\xi_n) \\ &\qquad \leqslant nP\big(|X_{\theta_n}| = a_n(\theta_n)\big) + nE|X_{\xi_n}|I_{(|X_{\xi_n}|>a_n(\xi_n))}/a_n(\xi_n) \\ &\qquad \to 0 \end{aligned}$$

by (2.6) and (2.7). From this fact, (2.13) also follows.

3. Central limit theorem in $\mathbb{R}^d$. Theorem 2 has an exact analogue in d-dimensions. A direction in $\mathbb{R}^d$ will be denoted by a unit vector $\vec{\phi} = (\phi_1, \cdots, \phi_d)$. If $Z = (X_1, \cdots, X_d)$ let $Z_{\vec{\phi}} = (Z,\vec{\phi}) \equiv \phi_1 X_1 + \cdots + \phi_d X_d$.

THEOREM 5. *Let* $Z_1, Z_2, \cdots$ *be* i.i.d. *copies of a purely d-dimensional random vector* $Z = (X_1, \cdots, X_d)$. *Assume* $EX_i = 0$ *for* $i = 1, \cdots, d$. *There exists a sequence of* $d \times d$ *matrices* A_n *such that*

$${}^t\big(A_n \Sigma_{i=1}^n {}^tZ_i\big) \to_{\mathfrak{D}} N(\vec{0}, I)$$

where I is the d-dimensional identity matrix, if and only if

$$\lim_{y\to\infty} \sup_{\{\vec{\phi}:\|\vec{\phi}\|=1\}} y^2 P\big(|X_{\vec{\phi}}| > y\big)/M_{\vec{\phi}}(y) = 0. \tag{3.1}$$

PROOF. The proof of necessity is analogous to that given in the 2-dimensional case so we sketch only the barest essentials. Let the $d \times d$ matrix A_n have the representation $A_n = (c_{i,j,n})$. Since ${}^t(A_n\Sigma_{i=1}^n {}^tZi) \to_{\mathfrak{D}} N(\vec{0}, I)$, the range of A_n is $\mathbb{R}^d$ for all n sufficiently large. Hence det $A_n \neq 0$ for all n sufficiently large. Thus,

$$Q_{i,n}^2 \equiv \left(\Sigma_{j=1}^d c_{i,j,n}^2\right)^{-1}$$

is finite so that

$$\vec{\theta}_{i,n} = Q_{i,n}(c_{i,1,n}, \cdots, c_{i,d,n})$$

is a unit vector in $\mathbb{R}^d$. These directions $\vec{\theta}_{i,n}$ are used to construct the d basis random variables $W_{\vec{\theta}_{i,n}}$ which are asymptotically i.i.d. $N(0, 1)$. Define

$$Z_{\vec{\theta}_{i,n},k} = Q_{i,n}\Sigma_{j=1}^d c_{i,j,n}X_{j,k}.$$

Then

$$\left(W_{\vec{\theta}_{1,n}}, \cdots, W_{\vec{\theta}_{d,n}}\right) \equiv \Sigma_{k=1}^n\left(Z_{\vec{\theta}_{1,n},k}/Q_{1,n}, \cdots, Z_{\vec{\theta}_{d,n},k}/Q_{d,n}\right)$$
$$= {}^t\left(A_n\Sigma_{k=1}^n {}^tZ_k\right).$$

The goal is to show that central limiting behavior occurs uniformly along each direction $\vec{\phi} \in \mathbb{R}^d$. Since det $A_n \neq 0$ for all n sufficiently large, for each unit vector $\vec{\phi} \in \mathbb{R}^d$ there exist unique constants $\beta_{1,n}(\vec{\phi}), \cdots, \beta_{d,n}(\vec{\phi})$ such that

$$Z_{\vec{\phi},k} = \beta_{1,n}(\vec{\phi})Z_{\vec{\theta}_{1,n},k}/Q_{1,n} + \cdots + \beta_{d,n}(\vec{\phi})Z_{\vec{\theta}_{d,n},k}/Q_{d,n}.$$

Applying Rao's lemma, for any sequence of unit vectors $\vec{\phi}_n \in \mathbb{R}^d$,

$$\mathcal{L}\left(\Sigma_{k=1}^n Z_{\vec{\phi}_n,k}/\left(\Sigma_{i=1}^d \beta_{i,n}^2(\vec{\phi}_n)\right)^{1/2}\right) \to_{\mathfrak{D}} N(0, 1) \qquad \text{as} \quad n \to \infty.$$

To complete "necessity", use Proposition 1 with J as the collection of unit vectors in $\mathbb{R}^d$ and proceed exactly as in the 2-dimensional case.

SUFFICIENCY. We must construct the n-stage coordinate system. Recall that $Z_{\vec{\phi}} = (Z, \vec{\phi})$ and $a_n^2(\vec{\phi}) = nM_{\vec{\phi}}(a_n(\vec{\phi}))$. There exists a unit vector $\vec{\gamma}_{1,n}$ such that

$$a_n(\vec{\gamma}_{1,n}) = \inf_{\{\vec{\phi}:\|\vec{\phi}\|=1\}} a_n(\vec{\phi}).$$

Let $\vec{\gamma}_{1,n} = (\gamma_{1,n}^{(1)}, \cdots, \gamma_{1,n}^{(d)})$. For any unit vector $\vec{\phi}$ such that

$$\left(\vec{\phi}, \vec{\gamma}_{1,n}\right) \equiv \Sigma_{i=1}^d \phi_i\gamma_{1,n}^{(i)} = 0,$$

$\vec{\gamma}_{1,n}$ is the minimal direction in the plane spanned by $\vec{\phi}$ and $\vec{\gamma}_{1,n}$. Furthermore, since $\vec{\phi}$ and $\vec{\gamma}_{1,n}$ are perpendicular,

$$EZ_{\vec{\gamma}_{1,n}}^{\prime}Z_{\vec{\phi}} = 0,$$

where for any $\vec{\theta}$, $Z_{\vec{\theta}}^{\prime} \equiv Z_{\vec{\theta}}I_{(|Z_{\vec{\theta}}| \leq a_n(\vec{\theta}))}$.

Having defined $\vec{\gamma}_{1,n}, \cdots, \vec{\gamma}_{k-1,n}$ where $k - 1 < d$, let $\vec{\gamma}_{k,n}$ be a unit vector in $\Gamma_{k-1}^{\perp}$ satisfying

$$a_n(\vec{\gamma}_{k,n}) = \inf_{\vec{\phi}\in\Gamma_{k-1}^{\perp}} a_n(\vec{\phi})$$

where $\Gamma_{k-1}^{\perp} = \{\vec{\phi} \in \mathbb{R}^d: \|\vec{\phi}\| = 1 \text{ and } (\vec{\phi}, \vec{\gamma}_{i,n}) = 0 \text{ for } i = 1, 2, \cdots, k - 1\}$. Thus, $\vec{\gamma}_{k,n}$ is the "nth minimal direction" in the orthogonal complement in $\mathbb{R}_d$ of the subspace spanned by $\{\vec{\gamma}_{1,n}, \cdots, \vec{\gamma}_{k-1,n}\}$.

Take $1 \leqslant i < j \leqslant d$. Clearly, $\vec{\gamma}_{i,n}$ is the "nth minimal direction" in the plane spanned by $\vec{\gamma}_{i,n}$ and $\vec{\gamma}_{j,n}$. Hence $EZ_{\vec{\gamma}_{i,n}}^{\perp} Z_{\vec{\gamma}_{j,n}} = 0$. Consequently, the truncated covariances are "asymptotically uncorrelated" in the sense analogous to (2.13). Then using exactly the same analysis as in Theorem 2, one may prove that

$$\mathcal{L}(U_{1,n}/a_n(\vec{\gamma}_{1,n}), \cdots, U_{d,n}/a_n(\vec{\gamma}_{d,n})) \to_{\mathfrak{D}} N(\vec{0}, I)$$

where $U_{i,n} = \sum_{k=1}^{n}(Z_{\vec{\gamma}_{i,n}})_k$. □

4. Determination of norming constants for X_θ. Theorem 2 indicates how to determine norming constants which yield one-dimensional central limiting behavior of X_θ. These are defined in terms of the norming constants of two asymptotically uncorrelated, time varying directions. It is natural to inquire when fixed directions which are not necessarily uncorrelated can be used. The next proposition supplies a sufficient condition.

PROPOSITION 6. *Suppose X_ϕ and $X_{\phi+\pi/2}$ both satisfy the one-dimensional* CLT *with norming constants a_n and b_n respectively. Define*

$$\tau_n = nEX_\phi X_{\phi+\pi/2} I_{(|X_\phi| < a_n, |X_{\phi+\pi/2}| < b_n)}$$

and $r_n^2(\theta) = a_n^2 \cos^2(\theta - \phi) + b_n^2 \sin^2(\theta - \phi) + \tau_n \sin 2(\theta - \phi)$. If there exists $\eta > 0$ and N such that for $n \geqslant N$,

$$|\tau_n|/a_n b_n < 1 - \eta, \tag{4.1}$$

then each X_θ satisfies the one-dimensional CLT *and $r_n(\theta)$ is an appropriate sequence of norming constants.*

PROOF. Without loss of generality, we assume $\phi = 0$ so $(X_\phi, X_{\phi+\pi/2}) = (X, Y)$. Let $c_n^2(\theta) = a_n^2 \cos^2 \theta + b_n^2 \sin^2 \theta$. By (4.1),

$$r_n^2(\theta)/c_n^2(\theta) = 1 + \tau_n \sin 2\theta / c_n^2(\theta) \geqslant 1 - |\tau_n|/a_n b_n \geqslant \eta.$$

Since X and Y both satisfy the CLT, given $\varepsilon > 0$,

$$\begin{aligned} nP(|X_\theta| > \varepsilon r_n(\theta)) &\leqslant nP(|X \cos \theta| > \varepsilon r_n(\theta)/2) + nP(|Y \sin \theta| > \varepsilon r_n(\theta)/2) \\ &\leqslant nP\big(|X \cos \theta| > \varepsilon \eta^{\frac{1}{2}} c_n/2\big) + nP\big(|Y \sin \theta| > \varepsilon \eta^{\frac{1}{2}} c_n/2\big) \\ &\leqslant nP\big(|X| > \varepsilon \eta^{\frac{1}{2}} a_n/2\big) + nP\big(|Y| > \varepsilon \eta^{\frac{1}{2}} b_n/2\big) \to 0 \quad \text{as} \quad n \to \infty. \end{aligned} \tag{4.2}$$

Moreover,

$$\begin{aligned} &nEX_\theta^2 I_{(|X_\theta| > r_n(\theta), |X| < a_n, |Y| < b_n)}/r_n^2(\theta) \\ &\leqslant 2(c_n(\theta)/r_n(\theta))^2 nP(|X_\theta| > r_n(\theta)) \to 0 \quad \text{as} \quad n \to \infty. \end{aligned} \tag{4.3}$$

Hence, using "$\sim$" to denote "is asymptotic to,"

$$(4.4) \qquad nL_\theta(r_n(\theta))/r_n^2(\theta) \sim \frac{n}{r_n^2(\theta)} EX_\theta^2 I_{(|X_\theta|<r_n(\theta),\, |X|<a_n, |Y|<b_n)}$$
$$\sim nEX_\theta^2 I_{(|X|<a_n, |Y|<b_n)}/r_n^2(\theta).$$

Expanding X_θ^2 and using the fact that

$$nE\left(X^2 \cos^2 \theta I_{(|X|<a_n, |Y|>b_n)} + Y^2 \sin^2 \theta I_{(|X|>a_n, |Y|<b_n)}\right)/c_n^2(\theta)$$
$$\leqslant nP(|Y| > b_n) + nP(|X| > a_n) \to 0,$$

(4.4) is asymptotic to

$$\left[nL_0(a_n)\cos^2 \theta + nL_{\pi/2}(b_n)\sin^2 \theta + \tau_n \sin 2\theta\right]/r_n^2(\theta)$$
$$\sim \left(a_n^2 \cos^2 \theta + b_n^2 \sin^2 \theta + \tau_n \sin 2\theta\right)/r_n^2(\theta) = 1.$$

So the $r_n(\theta)$ are indeed appropriate norming constants.

We show that X_θ satisfies the CLT by verifying condition 1.3. Because X_ϕ and $X_{\phi+\pi/2}$ satisfy the 1-dimensional CLT, $a_{2n}^2/a_n^2 \sim 2 \sim b_{2n}^2/b_n^2$. Moreover, $\tau_{2n}/\tau_n \sim 2$. Therefore, $\lim_{n\to\infty} r_{2n}^2(\theta)/r_n^2(\theta) = 2$. Now if $r_n(\theta) \leqslant y \leqslant r_{2n}(\theta)$, again utilizing (4.2) and (4.4),

$$y^2P(|X_\theta| > y)/M_\theta(y) \leqslant y^2P(|X_\theta| > y)/L_\theta(y)$$
$$\leqslant r_{2n}^2(\theta)P(|X_\theta| > r_n(\theta))/L_\theta(r_n(\theta))$$
$$= \left(r_{2n}(\theta)/r_n(\theta)\right)^2 r_n^2(\theta)P(|X_\theta| > r_n(\theta))/L_\theta(r_n(\theta))$$
$$\sim 2nP(|X_\theta| > r_n(\theta)) \to 0 \text{ as } n \to \infty.$$

Hence (1.3) holds. □

Note that $c_n(\theta)$ is asymptotic to $r_n(\theta)$ and so is an appropriate norming constant for X_θ when $\lim_{n\to\infty} \tau_n/a_nb_n = 0$. This holds when X_ϕ and $X_{\phi+\pi/2}$ are independent and more generally when $\phi = \gamma_n$.

5. Examples. To shed further light on the main theorem, we present a number of examples. The first illustrates one way in which componentwise normalization may fail even though matrix normalization succeeds. It also shows that the norming constants for X_θ cannot always be computed by calculating $r_n(\theta)$.

EXAMPLE 1. Let (X, Y) be a rotation of the independent variables (U, V) belonging to the domain of attraction of the normal distribution with norming constants a_n and b_n respectively;

$$(X, Y) = (U \cos \phi - V \sin \phi,\ U \sin \phi + V \cos \phi).$$

If $a_n/b_n \to 0$ as $n \to \infty$, then it is clear without any calculations that one has to rotate back ϕ (mod $\pi/2$) and normalize componentwise in order to get a nondegenerate limiting distribution.

Furthermore, $X_\theta = U\cos(\theta - \phi) + V\sin(\theta - \phi)$ and hence $a_n(\theta) \sim |\sin(\theta - \phi)|b_n$ if $\theta \neq \phi(\bmod \pi)$ and $a_n(\phi) \sim a_n$. So if $\phi \neq 0(\bmod \pi/2)$,

$$\begin{aligned} r_n^2(\phi) &\sim nEX_\phi^2 I_{(|X|<a_n(0),|Y|<a_n(\pi/2))} \\ &\sim nEX_\phi^2 I_{(|X_\phi|<b_n)} \\ &= nL_\phi(b_n). \end{aligned}$$

Consequently, $r_n^2(\phi)/a_n^2(\phi) \sim L_\phi(b_n)/L_\phi(a_n) \to \infty$ provided L_ϕ is properly chosen. $\square$

The second example illustrates that componentwise normalization may also fail because the truncated correlations $\rho_n(\theta, \theta') \equiv nEX'_\theta X'_{\theta'}/a_n(\theta)a_n(\theta')$ vary without tending to a limit as n increases. To obtain fixed correlations here, a varying coordinate system is required. An amusing aspect of this example is that although Y is strictly functionally dependent on X, nevertheless, along an appropriate subsequence $j_{2k}, \Sigma_{j=1}^{j_{2k}} Y_j$ is asymptotically independent of $\Sigma_{j=1}^{j_{2k}} X_j$.

EXAMPLE 2. Let X be a symmetric random variable with infinite variance such that $\lim_{y\to\infty} y^2 P(|X| > y) = 0$. Let $Y = XI_{(|X| \in A)}$, where $A = \cup_{n=1}^{\infty}(c_{2n}, c_{2n+1})$ and c_n satisfies $M_0(c_n) \equiv E(X^2 \wedge c_n^2) = n!$.

The pair (X, Y) obeys the CLT in $\mathbb{R}^2$ since

$$\begin{aligned} \sup_\theta y^2 P(|X_\theta| > y) &\leq y^2(P(|X| > y/2) + P(|Y| > y/2)) \\ &\leq 2y^2 P(|X| > y/2) \to 0 \qquad \text{as } y \to \infty \end{aligned}$$

so that condition (2.1) is satisfied.

Suppose for the moment that componentwise normalization gives weak convergence to $N(\vec{0}, I)$. Then

$$\begin{aligned} 0 = \lim_{n\to\infty} nEX'Y'/a_n b_n &= \lim_{n\to\infty} nE(Y')^2/a_n b_n \\ &= \lim_{n\to\infty} b_n/a_n, \end{aligned}$$

where $\{a_n\}$ and $\{b_n\}$ are the appropriate norming constants for X and Y, respectively, and $X' = XI(|X| \leq a_n)$, $Y' = YI(|Y| \leq b_n)$. Select j_k such that $a_{j_k} \sim c_k$. For k even, $b_{j_k} = o(a_{j_k})$ and the truncated pairwise correlations do tend to zero. However, for k odd, $b_{j_k} \sim a_{j_k}$ and $j_k EX'Y'/a_{j_k} b_{j_k} \to 1 \neq 0$, which gives a contradiction. Not too surprisingly, for $\theta = -\pi/4$ and n tending to infinity along $\{j_{2k+1}\}$, $\cos\theta \Sigma_{i=1}^n X_i/a_n + \sin\theta \Sigma_{i=1}^n Y_i/b_n$ tends to zero rather than $N(0, 1)$ in distribution. $\square$

The next example shows that X_θ satisfying the CLT for all $\theta \in [0, \pi) \sim \{\phi\}$ does not imply the CLT for X_ϕ. Moreover, it illustrates that a sum $W + Z$ of independent random variables, W and Z, being in the domain of attraction of a normal does not guarantee that both W and Z are in the domain, contrary to a conjecture of H. Tucker.

EXAMPLE 3. Let W and Z be independent symmetric random variables such that $M_W(t)$ is slowly varying, $M_Z(t)$ is not, and $M_Z(t)/M_W(t) \to 0$ as $t \to \infty$. If a_n

are the norming constants for W, then $\mathcal{L}(a_n^{-1}\Sigma_1^n W_i) \to N(0, 1)$, but $\mathcal{L}(a_n^{-1}\Sigma_1^n Z_i) \to \delta_0$ because (by Feller, page 235),

$$P(|\Sigma_1^n Z_i| > \varepsilon a_n) \leqslant \frac{nL_z(\varepsilon a_n)}{\varepsilon^2 a_n^2} + nP(|Z| > \varepsilon a_n)$$

$$= \frac{nM_Z(\varepsilon a_n)}{\varepsilon^2 a_n^2} \sim \frac{M_Z(\varepsilon a_n)}{\varepsilon^2 M_W(\varepsilon a_n)} \to 0 \qquad \text{as} \quad n \to \infty.$$

So with appropriately chosen Z and W we have

$$X_\theta = Z\cos(\theta - \phi) + W\sin(\theta - \phi)$$

and hence X_θ belongs to the domain of attraction of the normal distribution except when $\theta = \phi (\text{mod } \pi)$.

For a specific example let

$$P(|W| > t) = t^{-2}, \qquad t \geqslant 1$$

$$P(|Z| = e^{2^n}) = c2^{n/2}/e^{2^{n+1}} \qquad \text{for} \quad n = 1, 2, \cdots. \qquad \square$$

One might wonder whether or not the uniformity condition, (2.1), is equivalent to

(5.1) $$\lim_{n\to\infty} y^2 P(|X_\theta| > y)/M_\theta(y) = 0$$

for each $\theta \in [0, 2\pi]$. Somewhat surprisingly, this is weaker than (2.1). By a modification of Example 3 we obtain a pair (X, Y) which does not satisfy (2.1) even though each X_θ satisfies (5.1), and hence the CLT. Such a situation can never occur for random vectors with finite second moments.

EXAMPLE 4. As in Example 3, select symmetric random variables W and Z such that $M_W(t)$ is slowly varying, $M_Z(t)$ is not, and $M_Z(t)/M_W(t) \to 0$. To construct X_θ we slowly rotate the coordinate system, picking off various sections of W and Z. Since $M_Z(t)$ is not slowly varying, there exists $\varepsilon > 0$ such that $\limsup_{t\to\infty} t^2 P(t < |Z| \leqslant 2t)/M_Z(t) > 2\varepsilon$. Choose $0 = t_0 < t_1 < t_2 < \cdots \nearrow \infty$ such that

(5.2) $$M_W(t_{k-1}) = o(L_Z(t_k/2))$$

and

(5.3) $$t_k^2 P(t_k/2 < |Z| \leqslant t_k) \geqslant 8\varepsilon M_Z(t_k/2).$$

Now set

(5.4) $$X_\theta = \Sigma_{k=1}^\infty(\cos(\theta - \theta_k)) Z I_{(t_{k-1} < |Z| < t_k)} + \Sigma_{k=1}^\infty(\sin(\theta - \theta_k)) W I_{(t_{k-1} < |W| < t_k)},$$

where $1 \geqslant \theta_1 > \theta_2 > \cdots$ tends to zero slowly enough to satisfy

(5.5) $$\frac{\theta_{k+1}}{\theta_k} \to 1,$$

(5.6) $$\lim_{k\to\infty} \sup_{t_{k-1} \leqslant t \leqslant t_k} M_Z(t)/\theta_k^2 M_W(t) = 0,$$

and

(5.7) $$\lim_{k\to\infty} \sup_{t_{k-1}\leqslant t\leqslant t_k} M_W(t/\theta_k)/M_W(t) = 1.$$

By a slight abuse of notation, write $X_\theta = Z_\theta + W_\theta$. For fixed θ with $0 < \theta < \pi$, $L_{W_\theta}(t) \sim \sin^2\theta L_W(t) \sim \sin^2\theta M_W(t)$ as $t \to \infty$. This is a slowly varying function of t. For $\theta = 0$ and $t_{k-1} \leqslant t \leqslant t_k$, $L_{W_0}(t) \sim \theta_k^2 L_W(t/\theta_k) \sim \theta_k^2 M_W(t/\theta_k)$. By (5.7) this in turn is asymptotic to $\theta_k^2 M_W(t)$, which by (5.5) is also slowly varying as $t \to \infty$. Observe that for fixed θ,

$$\limsup_{t\to\infty} \frac{M_{Z_\theta}(t)}{M_{W_\theta}(t)} \leqslant \limsup_{t\to\infty} \frac{M_Z(t)}{M_{W_\theta}(t)}$$

$$\leqslant \limsup_{k\to\infty} \sup_{t_{k-1}\leqslant t\leqslant t_k} \frac{M_Z(t)}{\theta_k^2 M_W(t)} = 0 \qquad \text{by (5.6).}$$

Therefore, by the independence of Z_θ and W_θ, it is not difficult to show that $L_{X_\theta}(t) \sim L_{W_\theta}(t)$. Thus, for each fixed θ, $L_{X_\theta}(t)$ is slowly varying. This implies that X_θ is in the domain of attraction of the normal law and

$$\lim_{t\to\infty} \frac{t^2 P(|X_\theta| > t)}{L_\theta(t)} = 0.$$

(Here $L_\theta(t) \equiv L_{X_\theta}(t)$.) The convergence, however is not uniform in θ, since

$$\limsup_{t\to\infty} \sup_\theta \frac{t^2 P(|X_\theta| > t)}{L_\theta(t)} \geqslant \limsup_{k\to\infty} \frac{t_k^2 P(|X_{\theta_k}| > t_k/2)}{4L_{\theta_k}(t_k/2)}$$

$$\geqslant \limsup_{k\to\infty} \frac{t_k^2 P(t_k/2 < |Z| \leqslant t_k,\ W_{\theta_k} Z \geqslant 0)}{4M_Z(t_k/2)}$$

by (5.2)and the independence of Z and W

$$\geqslant \limsup_{k\to\infty} \frac{t_k^2 P(t_k/2 < |Z| \leqslant t_k)}{8M_Z(t_k/2)} \qquad \text{by symmetry}$$

$$\geqslant \varepsilon > 0 \qquad \text{by(5.3).}$$

Consequently, the pair $(X_0, X_{\pi/2})$ does not generate matrix normalized sums converging to a two-dimensional normal distribution.

An example of such an X_θ is afforded by taking the specific W and Z of Example 3, letting $\theta_k = 1/k$ and $t_k = \exp 2^{4^k}$.

Remark added in proof. Hahn and Klass, in (1980a) and (1980b), respectively, derive an extension of Theorem 2 in the nonidentically distributed case and an analogue for spherically symmetric stable limits in the i.i.d. case. Resnick and

Greenwood (1979) consider stable limit distributions in $\mathbb{R}^2$ using componentwise norming. Using linear operators, Kandelaki and Sazonov (1964) give a Hilbert space analogue of the Lindeberg-Fuller theorem when the random elements have finite second moments. As discussed in Hahn (1979), the infinite variance case in Hilbert space presents new pathologies. Sharpe (1969) characterized all possible limit laws in $\mathbb{R}^d$ (called operator stable laws) arising from affinely transformed partial sums of i.i.d. random vectors. Urbanik and others have pursued the study of operator stable laws.

Acknowledgment. We would like to express our gratitude to the referee, whose suggestions led to substantial simplification in the exposition of the examples, and to Lucien LeCam, who provided valuable discussion.

REFERENCES

[1] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

[2] FELLER, W. (1966). *An Introduction to Probability Theory and its Applications, Vol. II.* Wiley, New York.

[3] HAHN, M. G. (1979). The generalized domain of attraction of a Gaussian law on Hilbert space. *Lecture Notes in Math.* **707** 125–144.

[4] HAHN, M. G. and KLASS, M. J. (1980a). The multidimensional central limit theorem for arrays normed by affine transformations. To appear *Ann. Probability*.

[5] HAHN, M. G. and KLASS, J. J. (1980b). The generalized domain of attraction of spherically symmetric stable laws on $\mathbb{R}^d$. To appear in *Proc. Conf. Probability on Vector spaces* II, Poland, *Lecture Notes in Math.*

[6] KANDELAKI, N. N. and SAZONOV, V. V. (1964). On the central limit theorem for random elements with values in Hilbert space. *Theor. Probability Appl.* **9** 38–46.

[7] LOÈVE, M. (1963). *Probability Theory*. Van Nostrand.

[8] RESNICK, S. and GREENWOOD, P. (1979). A bivariate stable characterization and domains of attraction. *J. Multivariate Anal.* **9** 206–221.

[9] SHARPE, M. (1969). Operator stable probability distributions on vector groups. *Trans. Amer. Math. Soc.* **136** 51–65.

DEPARTMENT OF MATHEMATICS
TUFTS UNIVERSITY
MEDFORD, MASS. 02155

DEPARTMENTS OF MATHEMATICS AND STATISTICS
UNIVERSITY OF CALIFORNIA AT BERKELEY
BERKELEY, CA 94720