Journal of Theoretical Probability, Vol. 12, No. 3, 1999

Sample Covariance Matrix for Random Vectors with Heavy Tails[1]

Mark M. Meerschaert[2] and Hans-Peter Scheffler[3]

Received January 15, 1998; revised October 9, 1998

We compute the asymptotic distribution of the sample covariance matrix for independent and identically distributed random vectors with regularly varying tails. If the tails of the random vectors are sufficiently heavy so that the fourth moments do not exist, then the sample covariance matrix is asymptotically operator stable as a random element of the vector space of symmetric matrices.

KEY WORDS: Operator stable; generalized domains of attraction; regular variation; sample covariance matrix; heavy tails.

1. INTRODUCTION

Suppose that $X, X_1, X_2, X_3,...$ are independent, identically distributed random vectors on $\mathbb{R}^d$. In this paper we compute the asymptotic distribution of the sample covariance matrix

$$M_n = \sum_{i=1}^{n} X_i X_i' \tag{1.1}$$

when the distribution of X has heavy tails. When $E\|X\|^4 < \infty$ it is well known that M_n is asymptotically normal. In this paper we show that, when X has infinite fourth moments and the distribution of X is regularly varying, M_n is asymptotically operator stable as a random element of the vector space of symmetric $d \times d$ real matrices. The limiting distribution is

---

[1] This research was supported by a research scholarship for the Volkswagen Stiftung Research in Pairs program at Oberwolfach, Germany.

[2] Department of Mathematics, University of Nevada, Reno, Nevada 89557. E-mail: mcubed@unr.edu.

[3] Department of Mathematics, University of Dortmund, 44221 Dortmund Germany. E-mail: hps@mathematik.uni-dortmund.de.

0894-9840/99/0700-0821$16.00/0

not necessarily full dimensional on this space, however we prove that the limit is almost surely invertible. We conclude this paper with an application to self-normalized sums, where we answer a question posed by Vu *et al.*[18]

Regular variation[3] is an analytic growth condition used to provide elegant necessary and sufficient conditions for the central limit theorem to hold, see Feller[7] XVII.5 for the one variable case. A sequence of linear operators on $\mathbb{R}^d$ is regularly varying with index $(-E)$ if

$$A_{[tn]} A_n^{-1} \to t^{-E} \tag{1.2}$$

for all $t > 0$. Here $t^{-E} = \exp(-E \log t)$ where $\exp(A) = I + A + A^2/2! + A^3/3! + \cdots$ is the usual exponential operator. Let $\mathscr{B}$ denote the set of σ-finite Borel measures on $\Gamma = \mathbb{R}^k - \{0\}$. Topologize by writing $\mu_n \to \mu$ if and only if $\mu_n(S) \to \mu(S)$ for relatively compact Borel subsets of Γ whose topological boundary has μ-measure zero. Define $A\mu(S) = \mu(A^{-1}S)$. We say that $\mu \in \mathscr{B}$ varies regularly with index E if

$$nA_n \mu \to \phi \tag{1.3}$$

where A_n varies regularly with index $(-E)$ and $\phi \in \mathscr{B}$ cannot be supported on any $d-1$ dimensional subspace of $\mathbb{R}^d$. It follows from the regular variation of the sequence A_n that

$$t\phi(dx) = (t^E \phi)(dx) \tag{1.4}$$

for all $t > 0$, see Meerschaert.[10]

Operator stable laws[9] are the limiting distributions which appear in the central limit theorem for heavy tail random vectors. Suppose that Y is a random vector on $\mathbb{R}^d$ whose distribution is full, i.e., not supported on any $d-1$ dimensional affine subspace. If there exist linear operators A_n and nonrandom vectors a_n such that

$$A_n(X_1 + \cdots + X_n) - a_n \Rightarrow Y \tag{1.5}$$

then we say that X belongs to the generalized domain of attraction of Y. If $E\|X\|^2 < \infty$ then this is the usual central limit theorem with Y multivariate normal. Sharpe[16] shows that if (1.5) holds then Y is operator stable, meaning that if $Y, Y_1, Y_2, Y_3, \ldots$ are iid then for some linear operators L_n and nonrandom vectors b_n we have

$$L_n(Y_1 + \cdots + Y_n) - b_n \stackrel{d}{=} Y \tag{1.6}$$

for all n. Sharpe also shows that we can take $L_n = n^{-E}$ for some linear operator E called an exponent of the operator stable law, and that the limit Y decomposes into a normal component corresponding to the eigenvalues of E with real part equal 1/2, and a nonnormal component corresponding to the eigenvalues with real part exceeding 1/2. Meerschaert[12] shows that the norming operators A_n in (1.5) can always be chosen to vary regularly with exponent $(-E)$ where E is an exponent of the operators stable law. Meerschaert[11] shows that X belongs to generalized domain of attraction of a nonnormal operator stable law with exponent E if and only if its distribution varies regularly with exponent E. In other words (1.3) is a necessary and sufficient condition for (1.5) to hold with the same norming operator A_n, and the index of regular variation of μ equals the exponent of the nonnormal operator stable limit Y. It also turns out (see Theorem 1) that $A_n M_n A_n^* - B_n \Rightarrow W$ where A_n is the same as in (1.3) and B_n is a nonrandom centering. If every eigenvalue of E has real part exceeding 1/2, then Y is nonnormal and the same norming operators A_n are used for both the sum and the sample covariance. However the results of Theorem 1 are more general than that.

The exponent E governs the tail behavior and hence the moments of X. If every eigenvalue of E has real part exceeding 1/2 then X has infinite second moments, and the asymptotics of the sum in (1.5) are nonnormal. Asymptotics of the sample covariance (1.1) are dominated by squared terms, so in order to obtain normal asymptotics we need finite fourth moments. If the eigenvalues of E have real part exceeding 1/4 then the fourth moments of X are infinite, and the asymptotics of the sample covariance matrix will be nonnormal, the case considered in this paper. If the eigenvalues of E have real part between 1/4 and 1/2, then the sum has normal asymptotic's but the sample covariance matrix does not. In this case X has finite second moments, and the central limit theorem shows that (1.5) holds with $A_n = n^{-1/2} I$ instead of the norming operators from (1.3).

2. ASYMPTOTIC BEHAVIOR OF THE SAMPLE COVARIANCE MATRIX

In this section we compute the asymptotic distribution of the sample covariance matrix (1.1) when the distribution of X is regularly varying with infinite fourth moments. We begin by recalling a few facts about regular variation and limit theorems in $\mathbb{R}^d$. Define $\Lambda = \max\{\Re(\alpha)\}$ and $\lambda = \min\{\Re(\alpha)\}$ where α ranges over the eigenvalues of E. Meerschaert,[10] [Thm. 4.3] together with a uniform version of Seneta[14] [Thm. A.2] (see also Meerschaert,[11] Lemma 2) shows that if μ is regularly varying with exponent E then the moment functions

$$\begin{aligned} U_\zeta(r, \theta) &= E\,|\langle X, \theta\rangle|^\zeta I(|\langle X, \theta\rangle| \leqslant r) \\ V_\eta(r, \theta) &= E\,|\langle X, \theta\rangle|^\eta I(|\langle X, \theta\rangle| > r) \end{aligned} \tag{2.1}$$

are uniformly R—O varying whenever $\eta < 1/\Lambda \leqslant 1/\lambda < \zeta$, which means that for any $\delta > 0$ there exist real constants m, M, r_0 such that

$$\frac{V_\eta(tr, \theta)}{V_\eta(r, \theta)} \geqslant m t^{\eta - 1/\lambda - \delta}$$
$$\frac{U_\zeta(tr, \theta)}{U_\zeta(r, \theta)} \leqslant M t^{\zeta - 1/\Lambda + \delta} \tag{2.2}$$

for all $\|\theta\| = 1$, all $t \geqslant 1$ and all $r \geqslant r_0$. A uniform version of Feller[7] [p. 289] yields that for some positive real constants A, B, t_0 we have

$$A \leqslant \frac{t^{\zeta - \eta} V_\eta(t, \theta)}{U_\zeta(t, \theta)} \leqslant B \tag{2.3}$$

for all $\|\theta\| = 1$ and all $t \geqslant t_0$.

We will also employ standard results on infinitely divisible laws and convergence of triangular arrays, see Araujo and Giné[1] or Tortrat.[17] Suppose that Y, Y_1, Y_2, Y_3,... are independent, identically distributed random vectors on some finite dimensional real vector space V. We say that Y is infinitely divisible if for each n there exists a nonrandom vector $a_n \in V$ such that $Y_1 + \cdots + Y_n$ and $Y + a_n$ are identically distributed. The characteristic function of an infinitely divisible Y can be written uniquely in the form

$$Ee^{\langle t, Y \rangle} = \exp\left[i\langle t, a \rangle - \frac{1}{2} Q(t) + \int_{x \neq 0} e^{i\langle t, x \rangle} - 1 - \frac{i\langle t, x \rangle}{1 + \langle x, x \rangle} \phi(dx) \right] \tag{2.4}$$

where $a \in V$, $Q(t)$ is a nonnegative definite quadratic form on V, and ϕ is a σ-finite Borel measure on $V - \{0\}$ which satisfies

$$\int_{x \neq 0} \min\{\|x\|^2, 1\} \, \phi(dx) < \infty \tag{2.5}$$

The measure ϕ is called a Lévy measure and the triple $[a, Q, \phi]$ is called the Lévy representation of Y. If the distribution μ of X is regularly varying then the limit measure ϕ in (1.3) is also the Lévy measure of the limit Y in (1.5), see Meerschaert.[11] If X, X_1, X_2, X_3,... are independent random vectors on V with common distribution μ then the standard criteria for convergence of triangular arrays implies that (1.5) holds with Y having no normal component if and only if (1.3) holds for some Lévy measure ϕ along with

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \left\{ \int_{\|x\| \leqslant \varepsilon} |\langle x, t \rangle|^2 \, nA_n \mu(dx) - \left(\int_{\|x\| \leqslant \varepsilon} |\langle x, t \rangle| \, nA_n \mu(dx) \right)^2 \right\} = 0 \tag{2.6}$$

We can always choose

$$a_n = a + \int_{\|x\| \leqslant \varepsilon} x \, nA_n \mu(dx) \tag{2.7}$$

and then Y is infinitely divisible with Lévy representation $[a, 0, \phi]$.

Let $\mathscr{M}_s^d$ denote the vector space of $d \times d$ symmetric matrices with real entries together with the norm $\|A\| = \langle A, A \rangle^{1/2}$ where

$$\langle A, B \rangle = \sum_{i,j=1}^{d} A_{ij} B_{ij} \tag{2.8}$$

and A_{ij} the ij element of A. Define $T: \mathbb{R}^d \to \mathscr{M}_s^d$ by $Tx = xx'$ and for any linear operator A on $\mathbb{R}^d$ let $L_A(B) = ABA^*$ where A^* is the transpose of A. It is easy to check that $L_A^{-1} = L_{A^{-1}}$, $L_A(Tx) = T(Ax)$ and $\langle Tx, Ty \rangle = \langle x, y \rangle^2$. Since Tx is a polynomial in x it is open and continuous. Since the matrix Tx has rank one, the mapping T is not onto, but it is easy to check that the image of T spans $\mathscr{M}_s^d$. Since the sample covariance matrix

$$M_n = \sum_{i=1}^{n} X_i X_i' \tag{2.9}$$

is the sum of i.i.d. random elements of the vector space $\mathscr{M}_s^d$, we can apply the convergence criteria for triangular arrays of random vectors, along with regular variation, to compute the asymptotic distribution of M_n.

Theorem 1. Suppose that μ is regularly varying with exponent E and (1.3) holds. If every eigenvalue of E has real part exceeding 1/4 then

$$A_n M_n A_n^* - B_n \Rightarrow W \tag{2.10}$$

for some nonrandom B_n, where W is infinitely divisible on $\mathscr{M}_s^d$ with Lévy representation $[C, 0, T\phi]$.

Proof. Rewrite (2.10) in the form

$$L_{A_n}(TX_1 + \cdots + TX_n) - B_n \Rightarrow W \tag{2.11}$$

and note that $T\mu$ is the distribution of TX. In order to establish (2.11) we will apply the standard convergence criteria for triangular arrays on the vector space $\mathcal{M}_s^d$. First we show that $T\phi$ is a Lévy measure. Since $\|Tx\|^2 = \langle Tx, Tx \rangle = \langle x, x \rangle^2 = \|x\|^4$ we have

$$\int_{A \neq 0} \min\{\|A\|^2, 1\}\ T\phi(dA) = \int_{\|A\| \leqslant 1} \|A\|^2\ T\phi(dA) + T\phi\{A : \|A\| > 1\}$$

$$= \int_{\|x\| \leqslant 1} \|x\|^4\ \phi(dx) + \phi\{x : \|x\| > 1\}$$

Following Hirsch and Smale[8] [Chap. 6] we compute that for every $\delta > 0$ there exists a positive constant M such that $\|t^E\| \leqslant M t^{\lambda - \delta}$ for all $0 < t \leqslant 1$. Fix any $c > 1$ and let $Q = \{x : a \leqslant \|x\| \leqslant b\}$ where $0 < a < b$ are chosen so that the union of all $c^{-kE}(Q)$ for $k = 0, 1, 2,\ldots$ contains $\{x : 0 < \|x\| \leqslant 1\}$. Choose $\delta > 0$ such that $1/4 < \lambda - \delta < \lambda$, and apply (1.4) to get

$$\int_{0 < \|x\| \leqslant 1} \|x\|^4\ \phi(dx) \leqslant \sum_{k=0}^{\infty} \int_{c^{-kE}(Q)} \|x\|^4\ \phi(dx)$$

$$\leqslant \sum_{k=0}^{\infty} \int_{c^{-kE}(Q)} \|c^{-kE}\|^4\ \|c^{kE}x\|^4\ \phi(dx)$$

$$\leqslant M \sum_{k=0}^{\infty} c^{-4k(\lambda - \delta)} \int_Q \|x\|^4\ (c^{kE}\phi)(dx)$$

$$= M \sum_{k=0}^{\infty} c^{k(1 - 4(\lambda - \delta))} \int_Q \|x\|^4\ \phi(dx)$$

which is finite since $1 - 4(\lambda - \delta) < 0$. Since $\phi\{x : \|x\| > 1\} < \infty$ by assumption, we have shown that $T\phi$ is a Lévy measure. Since T is continuous, if S is a $T\phi$-continuity set then $T^{-1}S$ is a ϕ-continuity set. Then for all such sets we have $nL_{A_n}T\mu(S) = nTA_n\mu(S) = nA_n\mu(T^{-1}S) \to \phi(T^{-1}S) = T\phi(S)$.

We have shown that $nL_{A_n}T\mu \to T\phi$ where $T\phi$ is a Lévy measure. Now by (2.6) along with the Schwartz inequality, in order to establish (2.11) it will suffice to show that

$$\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \int_{\|A\| \leqslant \varepsilon} \langle A, B \rangle^2\ nL_{A_n}T_\mu(dA) = 0 \qquad (2.12)$$

for all unit vectors $B \in \mathcal{M}_s^d$. First suppose that $B = Tb$ for some $b \in \mathbb{R}^d$. Then

$$\begin{aligned}
\int_{\|A\| \leqslant \varepsilon} \langle A, B\rangle^2 \, nL_{A_n} T\mu(dA) &\leqslant \int_{|\langle A, B\rangle| \leqslant \varepsilon} \langle A, B\rangle^2 \, nL_{A_n} T\mu(dA) \\
&= \int_{|\langle A, L_{A_n^*} B\rangle| \leqslant \varepsilon} \langle A, L_{A_n^*} B\rangle^2 \, nT\mu(dA) \\
&= \int_{|\langle Tx, L_{A_n^*} B\rangle| \leqslant \varepsilon} \langle Tx, L_{A_n^*} B\rangle^2 \, n\mu(dx) \\
&= \int_{|\langle Tx, L_{A_n^*} Tb\rangle| \leqslant \varepsilon} \langle Tx, L_{A_n^*} Tb\rangle^2 \, n\mu(dx) \\
&= \int_{|\langle Tx, TA_n^* b\rangle| \leqslant \varepsilon} \langle Tx, TA_n^* b\rangle^2 \, n\mu(dx) \\
&= \int_{|\langle x, A_n^* b\rangle| \leqslant \varepsilon^{1/2}} \langle x, A_n^* b\rangle^4 \, n\mu(dx) \\
&= nr_n^4 U_4(\varepsilon^{1/2} r_n^{-1}, \theta_n)
\end{aligned}$$

where $r_n = \|A_n^* b\|$ and $\theta_n = A_n^* b / r_n$. Since $\lambda > 1/4$ we can apply (2.2) and (2.3) with $\eta = 0$ and $\zeta = 4$ to obtain

$$\begin{aligned}
nr_n^4 U_4(\varepsilon^{1/2} r_n^{-1}, \theta_n) &= nr_n^4 (\varepsilon^{1/2} r_n^{-1})^4 (\varepsilon^{1/2} r_n^{-1})^{-4} \, U_4(\varepsilon^{1/2} r_n^{-1}, \theta_n) \\
&= n\varepsilon^2 (\varepsilon^{1/2} r_n^{-1})^{-4} \, U_4(\varepsilon^{1/2} r_n^{-1}, \theta_n) \\
&\leqslant n\varepsilon^2 A^{-1} V_0(\varepsilon^{1/2} r_n^{-1}, \theta_n) \\
&\leqslant n\varepsilon^2 A^{-1} m^{-1} (\varepsilon^{1/2})^{-1/\lambda - \delta} \, V_0(r_n^{-1}, \theta_n) \\
&= \varepsilon^{2 - (1/2)(1/\lambda + \delta)} (Am)^{-1} \, nV_0(r_n^{-1}, \theta_n)
\end{aligned}$$

where we may choose $\delta < 4 - 1/\lambda$ so that $2 - (1/2)(1/\lambda + \delta) > 0$. Note also that

$$\begin{aligned}
nV_0(r_n^{-1}, \theta_n) &= n \int_{|\langle x, \theta_n\rangle| > r_n^{-1}} \mu(dx) \\
&= n \int_{|\langle x, r_n \theta_n\rangle| > 1} \mu(dx) \\
&= n \int_{|\langle x, A_n^* b\rangle| > 1} \mu(dx) \\
&= n\mu\{x : |\langle x, A_n^* b\rangle| > 1\} = n\mu\{x : |\langle A_n x, b\rangle| > 1\} \\
&= nA_n \mu\{x : |\langle x, b\rangle| > 1\} \to \phi\{x : |\langle x, b\rangle| > 1\}
\end{aligned}$$

and so (2.12) holds whenever $B = Tb$. Choose $\{b_k\}$ unit vectors such that $\{Tb_k\}$ spans $\mathscr{M}_s^d$. Note that $\|Tb\| = 1$ if and only if $\|b\| = 1$. Now for any $B \in \mathscr{M}_s^d$ we can write $B = \sum c_k Tb_k$ and then

$$\begin{aligned}
&\int_{\|A\| \leqslant \varepsilon} \langle A, B \rangle^2 \, nL_{A_n} T\mu(dA) \\
&= \int_{\|A\| \leqslant \varepsilon} \left\langle A, \sum c_k Tb_k \right\rangle^2 nL_{A_n} T\mu(dA) \\
&= \int_{\|A\| \leqslant \varepsilon} \left(\sum c_k \langle A, Tb_k \rangle \right)^2 nL_{A_n} T\mu(dA) \\
&= \sum_j \sum_k c_j c_k \int_{\|A\| \leqslant \varepsilon} \langle A, Tb_j \rangle \langle A, Tb_k \rangle \, nL_{A_n} T\mu(dA) \\
&\leqslant \sum_j \sum_k c_j c_k \left[\int_{\|A\| \leqslant \varepsilon} \langle A, Tb_j \rangle^2 \, nL_{A_n} T\mu(dA) \right. \\
&\qquad \left. \times \int_{\|A\| \leqslant \varepsilon} \langle A, Tb_k \rangle^2 \, nL_{A_n} T\mu(dA) \right]^{1/2} \\
&\leqslant \sum_j \sum_k c_j c_k \left[\int_{\langle A, Tb_j \rangle \leqslant \varepsilon} \langle A, Tb_j \rangle^2 \, nL_{A_n} T\mu(dA) \right. \\
&\qquad \left. \times \int_{\langle A, Tb_k \rangle \leqslant \varepsilon} \langle A, Tb_k \rangle^2 \, nL_{A_n} T\mu(dA) \right]^{1/2} \\
&= \sum_j \sum_k c_j c_k [nr_{nj}^4 U_4(\varepsilon^{1/2} r_{nj}^{-1}, \theta_{nj}) \, nr_{nk}^4 U_4(\varepsilon^{1/2} r_{nk}^{-1}, \theta_{nk})]^{1/2}
\end{aligned}$$

where $r_{nk} = \|A_n^* b_k\|$ and $\theta_{nk} = A_n^* b_k / r_{nk}$. Then as before we have

$$\begin{aligned}
&\sum_j \sum_k c_j c_k [nr_{nj}^4 U_4(\varepsilon^{1/2} r_{nj}^{-1}, \theta_{nj}) \, nr_{nk}^4 U_4(\varepsilon^{1/2} r_{nk}^{-1}, \theta_{nk})]^{1/2} \\
&\leqslant (Am)^{-1} \sum_j \sum_k c_j c_k \\
&\qquad \times [\varepsilon^{2-(1/2)(1/\lambda+\delta)} nV_0(r_{nj}^{-1}, \theta_{nj}) \, \varepsilon^{2-(1/2)(1/\lambda+\delta)} nV_0(r_{nk}^{-1}, \theta_{nk})]^{1/2} \\
&= (Am)^{-1} \varepsilon^{2-(1/2)(1/\lambda+\delta)} \sum_j \sum_k c_j c_k [nV_0(r_{nj}^{-1}, \theta_{nj}) \, nV_0(r_{nk}^{-1}, \theta_{nk})]^{1/2} \\
&\to (Am)^{-1} \varepsilon^{2-(1/2)(1/\lambda+\delta)} \sum_j \sum_k c_j c_k \\
&\qquad \times [\phi\{x : |\langle x, b_j \rangle| > 1\} \, \phi\{x : |\langle x, b_k \rangle| > 1\}]^{1/2}
\end{aligned}$$

as $n \to \infty$. Since $2-(1/2)(1/\lambda+\delta)>0$ we see that (2.12) holds, and so (2.11) holds as well.

Corollary 1. In Theorem 1, we can take $B_n=0$ when every eigenvalue of E has real part exceeding 1/2. We can take $B_n=nA_nEXX'A_n^*$ when every eigenvalue of E has real part less than 1/2.

Proof. Given a unit vector $b\in\mathbb{R}^d$ let $B=Tb$ and use (2.7) with $a=0$ to write

$$\begin{aligned}
|\langle B_n, Tb\rangle| &= \left|\left\langle \int_{\|A\|\leqslant R} AnL_{A_n}T\mu(dA), B\right\rangle\right| \\
&= \left|\int_{\|A\|\leqslant R} \langle A, B\rangle\, nL_{A_n}T\mu(dA)\right| \\
&\leqslant \int_{\|A\|\leqslant R} |\langle A, B\rangle|\, nL_{A_n}T\mu(dA) \\
&\leqslant \int_{|\langle A, B\rangle|\leqslant R} |\langle A, B\rangle|\, nL_{A_n}T\mu(dA) \\
&= nr_n^2 U_2(R^{1/2}r_n^{-1}, \theta_n)
\end{aligned}$$

where $r_n=\|A_n^*b\|$ and $\theta_n=A_n^*b/r_n$. If every eigenvalue of E has real part exceeding 1/2 then (2.3) applies with $\eta=0$ and $\zeta=2$ and so

$$\begin{aligned}
nr_n^2 U_2(R^{1/2}r_n^{-1}, \theta_n) &= nr_n^2(R^{1/2}r_n^{-1})^2\,(R^{1/2}r_n^{-1})^{-2}\,U_2(R^{1/2}r_n^{-1}, \theta_n) \\
&= nR(R^{1/2}r_n^{-1})^{-2}\,U_2(R^{1/2}r_n^{-1}, \theta_n) \\
&\leqslant A^{-1}RnV_0(R^{1/2}r_n^{-1}, \theta_n) \\
&\leqslant A^{-1}m^{-1}R^{1-(1/2)(1/\lambda+\delta)}nV_0(r_n^{-1}, \theta_n) \\
&\to A^{-1}m^{-1}R^{1-(1/2)(1/\lambda+\delta)}\phi\{x : |\langle x, b\rangle|>1\} \\
&\leqslant A^{-1}m^{-1}R^{1-(1/2)(1/\lambda+\delta)}\phi\{x : \|x\|>1\}
\end{aligned}$$

as $n\to\infty$. Then

$$\limsup_{n\to\infty} |\langle B_n, Tb\rangle| \leqslant CR^{\delta_0}$$

for $C=A^{-1}m^{-1}\phi\{x : \|x\|>1\}$ and $\delta_0=1-(1/2)(1/\lambda+\delta)>0$, for all unit vectors $b\in\mathbb{R}^d$. Taking $R>0$ arbitrarily small, we see that the convergence (2.10) still holds when $B_n=0$.

If every eigenvalue of E has real part less than 1/2 then (2.2) holds with $\zeta = 4$ and (2.3) holds with $\eta = 0$ and $\zeta = 4$, or $\eta = 2$ and $\zeta = 4$. Choose t_0, A, B so that both hold, and use (2.7) to write $B_n = nA_n EX_i X'_i A_n^* - I_n$ where

$$I_n = \int_{\|A\| > R} AnL_{A_n} T\mu(dA)$$

so that for any unit vector $b \in \mathbb{R}^d$ we have

$$\begin{aligned}
|\langle I_n, Tb\rangle| &= \left|\left\langle \int_{\|A\| > R} AnL_{A_n} T\mu(dA), Tb \right\rangle\right| \\
&\leqslant \int_{\|A\| > R} |\langle A, Tb\rangle| \, nL_{A_n} T\mu(dA) \\
&= \int_{|\langle A, Tb\rangle| > R} |\langle A, Tb\rangle| \, nL_{A_n} T\mu(dA) \\
&\quad + \int_{\|A\| > R \text{ and } |\langle A, Tb\rangle| \leqslant R} |\langle A, Tb\rangle| \, nL_{A_n} T\mu(dA)
\end{aligned}$$

where

$$\begin{aligned}
&\int_{|\langle A, Tb\rangle| > R} |\langle A, Tb\rangle| \, nL_{A_n} T\mu(dA) \\
&\quad = nr_n^2 V_2(R^{1/2} r_n^{-1}, \theta_n) \\
&\quad = nr_n^4 R^{-1} (R^{1/2} r_n^{-1})^2 \, V_2(R^{1/2} r_n^{-1}, \theta_n) \\
&\quad \leqslant nr_n^4 R^{-1} B U_4(R^{1/2} r_n^{-1}, \theta_n) \\
&\quad \leqslant MBR^{-1 + (1/2)(4 - 1/\Lambda + \delta)} nr_n^4 U_4(r_n^{-1}, \theta_n) \\
&\quad \leqslant MBA^{-1} R^{-1 + (1/2)(4 - 1/\Lambda + \delta)} n V_0(r_n^{-1}, \theta_n) \\
&\quad \to MBA^{-1} R^{-1 + (1/2)(4 - 1/\Lambda + \delta)} \phi\{x : |\langle x, b\rangle| > 1\} \\
&\quad \leqslant MBA^{-1} R^{-1 + (1/2)(4 - 1/\Lambda + \delta)} \phi\{x : \|x\| > 1\}
\end{aligned}$$

as $n \to \infty$, where $r_n = \|A_n^* b\|$ and $\theta_n = A_n^* b / r_n$ as before. Choose $\{b_k\}$ an orthonormal basis for $\mathbb{R}^d$ and note that if $\|y\|^2 > R$ then $\langle y, b_k \rangle^2 > R/d$ for some $k = 1,..., d$. Then

$$
\begin{aligned}
&\int_{\|A\| > R \text{ and } |\langle A, Tb \rangle| \leqslant R} |\langle A, Tb \rangle| \, n L_{A_n} T\mu(dA) \\
&\quad \leqslant R \int_{\|A\| > R} n L_{A_n} T\mu(dA) \\
&\quad = R \int_{\|A_n x\|^2 > R} n\mu(dx) \\
&\quad \leqslant R \sum_{k=1}^{d} \int_{\langle A_n x, b_k \rangle^2 > R/d} n\mu(dx) \\
&\quad = R \sum_{k=1}^{d} n V_0(r_{nk}^{-1} \sqrt{R/d}, \theta_{nk}) \\
&\quad \leqslant B d^2 R^{-1} \sum_{k=1}^{d} n r_{nk}^4 U_4(r_{nk}^{-1} \sqrt{R/d}, \theta_{nk}) \\
&\quad \leqslant M B d^{2-(1/2)(4-1/\Lambda+\delta)} R^{-1+(1/2)(4-1/\Lambda+\delta)} \sum_{k=1}^{d} n r_{nk}^4 U_4(r_{nk}^{-1}, \theta_{nk}) \\
&\quad = M B d^{2-(1/2)(4-1/\Lambda+\delta)} R^{-1+(1/2)(4-1/\Lambda+\delta)} \sum_{k=1}^{d} n \Lambda^{-1} V_0(r_{nk}^{-1}, \theta_{nk}) \\
&\quad \to M B d^{2-(1/2)(4-1/\Lambda+\delta)} R^{-1+(1/2)(4-1/\Lambda+\delta)} \\
&\qquad \times \sum_{k=1}^{d} \Lambda^{-1} \phi\{x : |\langle x, b_k \rangle| > 1\} \\
&\quad \leqslant M B d^{2-(1/2)(4-1/\Lambda+\delta)} R^{-1+(1/2)(4-1/\Lambda+\delta)} d \Lambda^{-1} \phi\{x : \|x\| > 1\}
\end{aligned}
$$

as $n \to \infty$, where $r_{nk} = \|A_n^* b_k\|$ and $\theta_{nk} = A_n^* b_k / r_{nk}$. Then

$$
\limsup_{n \to \infty} |\langle I_n, Tb \rangle| \leqslant C_1 R^{\delta_1}
$$

for all unit vectors $b \in \mathbb{R}^d$, where $\delta_1 = -1 + (1/2)(4 - 1/\Lambda + \delta) < 0$ for $\Lambda < 1/2$ and $C_1 = (B + B d^{3-(1/2)(4-1/\Lambda+\delta)}) M \Lambda^{-1} \phi\{x : \|x\| > 1\}$. Taking $R > 0$ arbitrarily large, we see that the convergence (2.10) still holds when $B_n = n A_n E X X' A_n^*$.

Remarks. Theorem 1 gives the asymptotic distribution of the uncentered sample covariance matrix M_n. In applications one often uses the centered sample covariance matrix

$$\Gamma_n = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)(X_i - \bar{X}_n)'$$

where $\bar{X}_n = n^{-1}(X_1 + \cdots + X_n)$ is the sample mean. Both M_n and $n\Gamma_n$ have the same asymptotics, see Meerschaert and Scheffler[13] where we prove this for the more general case of moving averages of i.i.d. random vectors. The asymptotic theory established in this paper is the foundation for the statistical applications appearing there. The regular variation condition in Theorem 1 is sufficient but not necessary. Regular variation includes a balancing condition on the tails of the measure μ which is stronger than necessary. For example in the one variable case we only need $P[|X| > x]$ regularly varying for the sample variance to be asymptotically stable, but we need an additional balanced tails condition for the sum to be asymptotically stable. Theorem 1 is new even in the special case where A_n are all diagonal. In this case Y is "marginally stable" meaning that the marginals of Y are all stable laws, with possibly different stable indices. When A_n are scalar multiples of the identity our result agrees with that of Davis *et al.*[5] and Davis and Marengo[6] but even here our approach provides additional information. In this case the limit W is multivariable stable with spectral measure TK where K is the spectral measure of the stable limit Y in (1.5), see the example at the end of Section 3 for an illustration.

3. THE LIMITING DISTRIBUTION

In this section we show that the limit W in Theorem 1 is operator stable, and we compute its exponent. We prove that W is almost surely invertible, but we provide an example to show that W need not have full dimension in the vector space $\mathcal{M}_s^d$.

Lemma 1. The limit W in (2.10) is operator stable with exponent ξ where $\xi M = EM + ME^*$.

Proof. Let $\psi = T\phi$ denote the Lévy measure of W. Since W has no normal component it suffices to show that $t\psi = t^{\xi}\psi$ for all $t > 0$. But $t\psi = tT\phi = Tt\phi = Tt^E\phi = L_{t^E}T\phi = L_{t^E}\psi$ for all $t > 0$. Since L_{t^E} is a one-parameter subgroup it can be written in the form $t^{\xi} = \exp(\xi \log t)$ for some element ξ

of the tangent space of $GL(\mathcal{M}_s^d)$. Then ξ is the derivative of $\exp(s\xi)$ at $s=0$, and since

$$\begin{aligned}\frac{\exp(s\xi)-I}{s}(M) &= \frac{L_{\exp(sE)}(M)-M}{s}\\ &= s^{-1}(\exp(sE)\,M\exp(sE^*)-M)\\ &= s^{-1}(\{I+sE+O(s^2)\}\;M\{I+sE^*+O(s^2)\}-M)\\ &= (EM+ME^*)+O(s)\end{aligned}$$

as $s\to 0$ we see that $\xi M = EM+ME^*$.

Theorem 2. If every eigenvalue of E has real part exceeding 1/2 then the limiting random matrix W in Theorem 1 is invertible with probability one.

Proof. Define $L=\text{span}\{\text{supp}(T\phi)\}$ so that L is a linear subspace of $\mathcal{M}_s^d$ and $T\phi$ is full on L. Since W is full and operator stable on L, a result of Sharpe shows that W has a density on L.

Define $\mathcal{M}_s^d(+)=\{M\in\mathcal{M}_s^d : \theta' M\theta\geqslant 0\ \ \forall\theta\in\mathbb{R}^d\}$ and let $L^+=L\cap\mathcal{M}_s^d(+)$. Since $T(X_i)\in\mathcal{M}_s^d(+)$ almost surely we have $M_n=\sum_{i=1}^n T(X_i)\in\mathcal{M}_s^d(+)$ almost surely and so $A_nM_nA_n^*\in\mathcal{M}_s^d(+)$ almost surely. Since $\mathcal{M}_s^d(+)$ is a closed subset of $\mathcal{M}_s^d$ we get using the Corollary of Theorem 1 that $P[W\in\mathcal{M}_s^d(+)]\geqslant\limsup_{n\to\infty}P[A_nM_nA_n^*\in\mathcal{M}_s^d(+)]$ by the Portmanteau theorem, and so $P[W\in\mathcal{M}_s^d(+)]=1$.

Since ϕ is full on $\mathbb{R}^d$ we can choose $\theta_1,\dots,\theta_d$ linearly independent in $\mathbb{R}^d$ with all $\theta_i\in\text{supp}(\phi)$. Otherwise $\text{span}\{\text{supp}(\phi)\}$ has dimension less than d and ϕ would not be full. Then $T\theta_i\in L^+$ for all $i=1,\dots,d$, and $T\theta_1,\dots,T\theta_d$ are linearly independent in L. If not then we can write $0=\sum_{i=1}^d c_iT\theta_i$ where not all $c_i=0$. Suppose for example that $c_1\neq 0$. Then $0=(\sum c_iT\theta_i)\,\theta_1=\sum c_i\langle\theta_i,\theta_1\rangle\,\theta_i$ but $c_1\langle\theta_1,\theta_1\rangle\neq 0$ which contradicts the linear independence of $\theta_1,\dots,\theta_d$.

Lemma 2. $\text{supp}(T\phi)=T\,\text{supp}(\phi)$.

Proof of Lemma 2. If $y\notin\text{supp}(T\phi)$ then there exists a $A\subset\mathcal{M}_s^d$ closed with $y\notin A$ and $T\phi(A^c)=0$. Let $B=T^{-1}A$. Since T is continuous B is closed, and $y\notin TB=TT^{-1}A=A$. Then $\phi(B^c)=\phi((T^{-1}A)^c)=\phi(T^{-1}(A^c))=T\phi(A^c)=0$, so $y\notin T\,\text{supp}(\phi)$. Conversely if $y\notin T\,\text{supp}(\phi)$ then for some $B\subset\mathcal{M}_s^d$ closed, $y\notin TB$ and $\phi(B^c)=0$. Let $A=TB$. Since T is an open mapping, A is closed and $y\notin A$. Then $B\subseteq T^{-1}TB=T^{-1}A$ so $T^{-1}(A^c)=(T^{-1}A)^c\subseteq B^c$. Then $T\phi(A^c)=\phi(T^{-1}(A^c))\leqslant\phi(B^c)=0$ so $y\notin\text{supp}(T\phi)$.

Lemma 2 implies that $L = \text{span}\{T\,\text{supp}(\phi)\}$, and so we can choose additional $\theta_i \in \text{supp}(\phi)$ as necessary so that $T\theta_1,\dots, T\theta_m$ form a basis for L. Note that $m \geqslant d$ and all $T\theta_i \in L^+$. Define $S = \{M \in L^+ : M \text{ singular}\}$ and let $M_0 = T\theta_1 + \cdots + T\theta_m$. For all nonzero $x \in \mathbb{R}^d$ we have $x'M_0 x = \sum_i x'T\theta_i x = \sum \langle x, \theta_i \rangle^2 > 0$ since $\theta_1,\dots, \theta_m$ span $\mathbb{R}^d$, so M_0 is positive definite. Additionally if $M \in L^+$ and $c > 0$ then $M + cM_0$ is positive definite because $x'(M + cM_0)\,x = x'Mx + cx'M_0 x \geqslant cx'M_0 x > 0$ whenever $x \neq 0$. Now define $V = \{M \in L : \langle M, M_0 \rangle = 0\}$, $V_M = \{M + tM_0 : t \in \mathbb{R}\}$, and $S_M = S \cap V_M$. If both $M + t_1 M_0$ and $M + t_2 M_0$ were in S_M then (assuming $t_2 > t_1$) $M + t_2 M_0 = (M + t_1 M_0) + (t_2 - t_1)\, M_0$ would be positive definite, and hence nonsingular, which is a contradiction. Then S_M can contain at most one point, and so S_M has Lebesgue measure zero in V_M for all $M \in V$. Then the Fubini theorem implies that S has Lebesgue measure zero in L. Since W has a density on L we have $P[W \in S] = 0$. Now $P[W \text{ singular}] = P[W \in S] + P[W \notin \mathcal{M}_s^d(+)] = 0 + 0 = 0$ which concludes the proof.

Example 1. The following example shows that the limit W in Theorem 1 is not necessarily full. Suppose we are on $\mathbb{R}^2$ and that the Lévy measure ϕ of the limit Y in (1.5) is concentrated on the positive coordinate axes with $\phi\{te_i : t > r\} = r^{-\alpha}$ for $i = 1, 2$ where $0 < \alpha < 2$. Then Y is full and operator stable on $\mathbb{R}^2$ with exponent $E = \alpha^{-1} I$, but W is not full. In particular if we write

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{and} \qquad Ty = \begin{pmatrix} y_1^2 & y_1 y_2 \\ y_1 y_2 & y_2^2 \end{pmatrix}$$

then $T\phi\{N : \langle M, N \rangle \neq 0\} = \phi\{y : \langle M, Ty \rangle \neq 0\} = \phi\{y : 2y_1 y_2 \neq 0\} = 0$ since ϕ is supported on the set $\{y : y_1 y_2 = 0\}$. In this case Y is actually multidimensional stable with index α and the spectral measure of Y is the sum of the unit masses at the points e_1, e_2. Then W is multivariable stable with index $\alpha/2$ on $\mathcal{M}_s^d$ and its spectral measure is the sum of the unit masses at the points $e_1 e_1'$, $e_2 e_2'$. Also $W \in L$ almost surely where $L = \text{span}\{\text{supp}(T\phi)\} = \text{span}\{e_1 e_1', e_2 e_2'\}$.

4. SELF-NORMALIZED SUMS

Assume that $X, X_1, X_2, X_3,\dots$ are i.i.d. on $\mathbb{R}^d$ with common distribution μ and that μ belongs to the strict generalized domain of attraction of some full operator stable law ν on $\mathbb{R}^d$ with no normal component. This means that

$$A_n S_n \Rightarrow Y \tag{4.1}$$

where A_n is some linear operator on $\mathbb{R}^d$, $S_n = \sum_{i=1}^n X_i$ and Y has distribution ν. Then μ varies regularly with exponent E where every eigenvalue of E has real part exceeding 1/2, and in fact $nA_n\mu \to \phi$ where ϕ is the Lévy measure of ν, and E is an exponent of ν. Let $M_n = \sum_{i=1}^n X_i X_i'$ denote the sample covariance matrix and apply Theorem 1 and Corollary 1 to see that

$$A_n M_n A_n^* \Rightarrow W \tag{4.2}$$

where W is an operator stable symmetric random matrix. By Theorem 2 we see that W is almost surely invertible. Then weak convergence implies that $P[A_n M_n A_n^* \text{ invertible}] \to 1$ as $n \to \infty$. Since Y is full, convergence of types[2] implies that A_n is invertible for all large n, and so the probability that M_n^{-1} exists tends to one as $n \to \infty$. To avoid complicated notation we will assume henceforth, without loss of generality, that M_n is invertible with probability one. We will write $M_n^{1/2}$ to denote the symmetric square root of the nonnegative definite symmetric matrix M_n.

Lemma 3. $A_n M_n^{1/2}$ is uniformly tight.

Proof. Note that for $A \in \mathscr{M}_s^d$ we have $\|A\|^2 = \text{trace}(AA^*)$. The continuous mapping theorem and (4.2) imply that $\text{trace}(A_n M_n A_n^*) \Rightarrow \text{trace}(W)$. Hence $\text{trace}(A_n M_n A_n^*)$ is uniformly tight. But then

$$\begin{aligned} P[\|A_n M_n^{1/2}\| > R] &= P[\|A_n M_n^{1/2}\|^2 > R^2] \\ &= P[\text{trace}(A_n M_n A_n^*) > R^2] < \varepsilon \end{aligned}$$

for all n and some $R > 0$.

Lemma 4. (a) $(A_n S_n, A_n M_n^{1/2})$ is uniformly tight; (b) $(A_n S_n, A_n M_n A_n^*)$ is uniformly tight.

Proof. For $x \in \mathbb{R}^d$ and $M \in \mathscr{M}_s^d$ define $\|(x, M)\| = \sqrt{\|x\|^2 + \|M\|^2}$. Then $P[\|(A_n S_n, A_n M_n^{1/2})\| > R] = P[\|A_n S_n\|^2 + \|A_n M_n^{1/2}\|^2 > R^2] \leqslant P[\|A_n S_n\|^2 > R^2/2] + P[\|A_n M_n^{1/2}\|^2 > R^2/2] < \varepsilon/2 + \varepsilon/2 = \varepsilon$ for all n and some $R > 0$ by (4.1) and Lemma 3. The proof of part (b) is similar.

Lemma 5. $(A_n S_n, A_n M_n A_n^*) \Rightarrow (Y, W)$.

Proof. Use the standard convergence criteria for triangular arrays. If U is a Borel subset of $\mathbb{R}^d - \{0\}$ and V is a Borel subset of $\mathscr{M}_s^d - \{0\}$ define $\Phi(U \times V) = \phi(U \cap T^{-1}V)$. The sets $U \times V$ form a convergence determining class. Now note that

$$\begin{aligned} nP[(A_nX, A_nXX'A_n^*)\in U\times V] &= nP[A_nX\in U, A_nXX'A_n^*\in V]\\ &= nP[A_nX\in U, T(A_nX)\in V]\\ &= nP[A_nX\in U, A_nX\in T^{-1}V]\\ &= nP[A_nX\in U\cap T^{-1}V]\\ &\to \phi(U\cap T^{-1}V)\\ &= \Phi(U\times V)\end{aligned}$$

so that $nL_n\lambda\to\Phi$ where λ is the distribution of (X, XX') on $\mathbb{R}^d\oplus\mathcal{M}_s^d$ and L_n is the linear operator on $\mathbb{R}^d\oplus\mathcal{M}_s^d$ defined by $L_n(x, M)=(A_nx, A_nMA_n^*)$. Note that the Lévy measure Φ of the joint limit is concentrated on the set $\{(x, xx'): x\in\mathbb{R}^d\}$. Now in order to establish joint convergence it suffices to show that

$$\lim_{\varepsilon\to 0}\limsup_{n\to\infty}\int_{\|(y,B)\|\leqslant\varepsilon}\langle(y,B),(x,A)\rangle^2\, nL_n\lambda\{dy, dB\}=0 \tag{4.3}$$

for all $(x, A)\in\mathbb{R}^d\oplus\mathcal{M}_s^d$. We use the inner product associated with the norm in Lemma 2, so that $R^d\perp\mathcal{M}_s^d$. Then

$$\begin{aligned}&\int_{\|(y,B)\|\leqslant\varepsilon}\langle(y,B),(x,A)\rangle^2\, nL_n\lambda\{dy, dB\}\\ &\quad=\int_{\|y\|^2+\|B\|^2\leqslant\varepsilon^2}\langle y,x\rangle^2+\langle B,A\rangle^2\, nL_n\lambda\{dy, dB\}\\ &\quad\leqslant\int_{\|y\|\leqslant\varepsilon}\langle y,x\rangle^2\, nA_n\mu\{dy\}+\int_{\|B\|\leqslant\varepsilon}\langle B,A\rangle^2\, nL_{A_n}T\mu\{dB\}\end{aligned}$$

and (4.3) follows from the convergence criteria for triangular arrays along with (1.5) and Theorem 1.

Theorem 3. If (4.1) and (4.2) hold then $M_n^{-1/2}S_n$ is weakly relatively compact with all limit points of the form $K^{-1}Y$ where KK^* and W are identically distributed.

Proof. By Lemma 4 part (a) there exists for every subsequence of the positive integers a further subsequence along which $(A_nS_n, A_nM_n^{1/2})\Rightarrow(Y, K)$ for some K. Since $(A_nM_n^{1/2})(A_nM_n^{1/2})^*=A_nM_nA_n^*\Rightarrow W$ while $A_nM_n^{1/2}\Rightarrow K$ we see that KK^* is identically distributed with W. Then K is almost surely invertible and the continuous mapping theorem implies that $M_n^{-1}S_n=(A_nM_n^{1/2})^{-1}(A_nS_n)\Rightarrow K^{-1}Y$ along this subsequence.

Theorem 4. If (4.1) and (4.2) hold with $A_n = a_n^{-1} I$ then $M_n^{-1/2} S_n \Rightarrow W^{-1/2} Y$.

Proof. In this case (4.1) becomes $a_n^{-1} S_n \Rightarrow Y$ and (4.2) reduces to $a_n^{-2} M_n \Rightarrow W$. Then Lemma 5 yields $(a_n^{-1} S_n, a_n^{-2} M_n) \Rightarrow (Y, W)$ where W is almost surely invertible by Theorem 2. Now the continuous mapping theorem implies that $M_n^{-1/2} S_n = (a_n^{-2} M_n)^{-1/2} (a_n^{-1} S_n) \Rightarrow W^{-1/2} Y$.

Remarks. When (1.5) holds with Y multivariable normal, Sepanski[15] shows that $M_n^{-1/2} S_n \Rightarrow Y$ even if $E\|X\|^2 = \infty$. Vu *et al.*[18] extend this result to dependent sequences, and they ask whether the same convergence holds for nonnormal limits. Theorem 4 show that this is not the case. We conjecture that we actually get weak convergence in Theorem 4 in the general case of operator norming, but we have not been able to prove this. Since the limiting distribution in Theorem 1 is not always full dimensional, one might ask whether a different norming would produce a full limit. However even when X has independent stable marginals their cross-product will not usually belong to any domain of attraction, see Cline.[4]

REFERENCES

1. Araujo, A., and Giné, (1980). *The Central Limit Theorem for Real and Banach Valued Random Variables*, Wiley, New York.
2. Billingsley, P. (1966). Convergence of types in k-space. *Z. Wahrsch. verw. Geb.* **5**, 175–179.
3. Bingham, N., Goldie, C., and Teugels, J. (1987). Regular Variation. *Encyclopedia of Mathematics and Its Applications*, Vol. 27, Cambridge University Press.
4. Cline, D. (1986). Convolution tails, product tails, and domains of attraction. *Prob. Th. Rel. Fields* **72**, 529–557.
5. Davis, R., Marengo, J., and Resnick, S. (1985). Extremal properties of a class of multivariate moving averages. *Proc. 45th Int. Statistical Institute*, Vol. **4**, Amsterdam.
6. Davis, R., and Marengo, J. (1990). Limit theory for the sample covariance and correlation matrix functions of a class of multivariate linear processes. *Commun. Statist. Stoch. Models* **6**, 483–497.
7. Feller, W. (1971). *An Introduction to Probability Theory and Its Applications*, Vol. II, Second Edition, Wiley, New York.
8. Hirsch, M., and Smale, S. (1974). *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press, New York.
9. Jurek, Z., and Mason, J. D. (1993). *Operator-Limit Distributions in Probability Theory*, Wiley, New York.
10. Meerschaert, M. (1988). Regular variation in $\mathbb{R}^k$, *Proc. Amer. Math. Soc.* **102**, 341–348.
11. Meerschaert, M. (1993). Regular variation and generalized domains of attraction in $\mathbb{R}^k$. *Stat. Prob. Lett.* **18**, 233–239.
12. Meerschaert, M. (1994) Norming operators for generalized domains of attraction. *J. Theor. Prob.* **7**, 793–798.

13. Meerschaert, M. and Scheffler, H. P. (1999). Moving averages of random vectors with regularly varying tails. *J. Time Series Anal.*, to appear.
14. Seneta, E. (1976). *Regularly Varying Functions. LNM*, Vol. 508, Springer, Berlin.
15. Sepanski, S. (1994). Asymptotics for multivariate t-statistic and Hotellings's T^2-statistic under infinite second moments via bootstrapping. *J. Multivariate Anal.* **49**, 41–54.
16. Sharpe, M. (1969). Operator-stable probability distributions on vector groups. *Trans. Amer. Math. Soc.* **136**, 51–65.
17. Tortrat, A. (1971). *Calcul des Probabilités et Introduction aux Processus Aléatoires* (French), Masson et Cie Éditeurs, Paris.
18. Vu, H., Maller, R., and Klass, M. (1996). On the studentization of random vectors. *J. Multivariate Anal.* **57**, 142–155.