DOI: 10.1214/ECP.v5-1020

Elect. Comm. in Probab. **5** (2000) 73–76

ELECTRONIC COMMUNICATIONS in PROBABILITY

A WEAK LAW OF LARGE NUMBERS FOR THE SAMPLE COVARIANCE MATRIX

STEVEN J. SEPANSKI
Department of Mathematics, Saginaw Valley State University, 7400 Bay Road University Center, MI 48710
email: `sepanski@svsu.edu`

ZHIDONG PAN
Department of Mathematics, Saginaw Valley State University, 7400 Bay Road University Center, MI 48710
email: `pan@svsu.edu`

submitted February 15, 1999 *Final version accepted* March 20, 2000

AMS 1991 Subject classification: Primary 60F05; secondary 62E20, 62H12
Law of large numbers, affine normalization, sample covariance, central limit theorem, domain of attraction, generalized domain of attraction, multivariate t-statistic

Abstract
In this article we consider the sample covariance matrix formed from a sequence of independent and identically distributed random vectors from the generalized domain of attraction of the multivariate normal law. We show that this sample covariance matrix, appropriately normalized by a nonrandom sequence of linear operators, converges in probability to the identity matrix.

1. Introduction:
Let $X, X_1, X_2 \cdots$ be iid R^d valued random vectors with $\mathcal{L}(X)$ full. The condition of fullness is the multivariate analogue of nondegeneracy and will be in force throughout this article. It means that $\mathcal{L}(X)$ is not concentrated on any $d-1$ dimensional hyperplane. Equivalently, $\langle X, \theta\rangle$ is nondegenerate for every θ. Here $\langle\ ,\ \rangle$ denotes the inner product.
Throughout this article all vectors in R^d are assumed to be column vectors. For any matrix, A, A^t denotes its transpose. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. We denote and define the sample covariance matrix by $C_n = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)(X_i - \bar{X}_n)^t$. That C_n has a unique nonnegative symmetric square root, denoted above by $C_n^{1/2}$, follows from the fact that $\langle C_n\theta, \theta\rangle = \sum_{i=1}^n \langle X_i - \bar{X}_n, \theta\rangle^2 \geq 0$, so that C_n is nonnegative. Also, C_n is clearly symmetric. However, there is no guarantee that C_n is invertible with probability one.
In [3] we describe two ways to circumvent the problem of lack of invertibility of C_n. One such approach is to define

$$B_n = \begin{cases} C_n & \text{if } C_n \text{ is invertible} \\ I & \text{otherwise} \end{cases} \tag{1.3}$$

The success of this approach relies on the fact that if $\mathcal{L}(X)$ is in the Generalized Domain of Attraction of the Normal Law (see (1.6) below for the definition), then $P(C_n = B_n) \to 1$. (See

[3], Lemma 5.) In light of this, we will assume without loss of generality that C_n is invertible. $\mathcal{L}(X)$ is said to be in the Generalized Domain of Attraction (GDOA) of the Normal Law if there exist matrices A_n and vectors v_n such that

$$A_n \sum_{i=1}^{n} X_i - v_n \Rightarrow N(0, I). \tag{1.6}$$

One construction of A_n is such that A_n is invertible, symmetric and diagonalizable. See Hahn and Klass [2].

The main result is Theorem 1 below. This result was shown in Sepanski [5]. However, there the proof was based on a highly technical comparison of the eigenvalues and eigenvectors of C_n and A_n. There the proof was essentially real valued. The purpose of this note is to give a more efficient proof that is operator theoretic and multivariate in nature. For more details, we refer the interested reader to the original article. In particular, Sepanski [5] contains a more complete list of references.

2. Results

Theorem 1: *If the law of X is in the generalized domain of attraction of the multivariate normal law, then*

$$\sqrt{n} A_n C_n^{1/2} \to I \quad \textit{in pr.}$$

Proof: Let $P_n(\omega)$ denote the empirical measure. That is, $P_n(\omega)(A) = \frac{1}{n} \sum_{i=1}^{n} I[X_i(\omega) \in A]$. Here I is the indicator function. For each $\omega \in \Omega$ let $X_1^*, \cdots X_n^*$ be iid with law $P_n(\omega)$.

Sepanski [4], Theorem 2, shows that under the hypothesis of GDOA,

$$A_n \sum_{j=1}^{n} X_j^* - n\mu \Rightarrow N(0, I) \quad \textit{in pr.}$$

Sepanski [3], Theorem 1, shows that under the hypothesis of GDOA,

$$(nC_n)^{-1/2} \sum_{j=1}^{n} X_j^* - n\mu \Rightarrow N(0, I) \quad \textit{in pr.}$$

These two results, together with the multivariate Convergence of Types theorem of Billingsley [1], imply that

$$(nC_n)^{-1/2} = B_n R_n A_n, \tag{1}$$

where $B_n \to I$ *in pr.*, and R_n are (random) orthogonal. The proof of Theorem 1 is thereby reduced to showing that $R_n \to I$ *in pr.* However, convergence in probability is equivalent to every subsequence having a further subsequence which converges almost surely. This reduces the proof to a pointwise result about the behavior of the linear operators.

Write $A_n = Q_n D_n Q_n^t$ where Q_n is orthogonal and D_n is diagonal with nonincreasing diagonal entries. Let $P_n = Q_n R_n Q_n^t$ and $K_n = Q_n B_n Q_n^t$.

$$\|K_n - I\| = \|Q_n^t B_n Q_n - Q_n^t Q_n\| \leq \|B_n - I\| \to 0$$

By the same token, $R_n \to I$ if and only if $P_n \to I$. Also, $(nC_n)^{-1/2}$ is positive and symmetric and therefore so are $B_n R_n A_n$ and $K_n P_n D_n$. The proof of Theorem 1 is reduced to the following lemma.

Lemma 2: Let P_n be orthogonal. Let $D_n = \text{diag}(\lambda_{n1}, \cdots, \lambda_{nd})$ be diagonal such that $\lambda_{n1} \geq \lambda_{n2} \geq \cdots \geq \lambda_{nd} > 0$. Suppose $K_n \to I$. If $K_n P_n D_n$ is positive and symmetric for every n, then $P_n \to I$.

Proof: Given a subsequence of P_n we show that there is a further subsequence along which $P_n \to I$. Let $E_n = \lambda_{n1}^{-1} D_n$. This is a diagonal matrix of all positive entries that are bounded above by 1. Therefore, given any subsequence, there is a further subsequence along which $K_n \to I$, $P_n \to P$, and $E_n \to E$. Necessarily, P is orthogonal and E is diagonal with entries in $[0, 1]$. Furthermore, E has at least one diagonal entry that is 1 and its entries are nonincreasing. Since $K_n P_n E_n$ is symmetric, nonnegative and $K_n \to I$, we have that $PE = EP^t$, and PE is nonnegative. Now, $(PE)^2 = (PE)^t PE = EP^{-1} PE = E^2$. Hence, since PE and E are both nonnegative, $PE = E$. If E is invertible, then $P = I$ and we are done. Suppose E is not invertible. Write $E = \begin{pmatrix} E_{(1)} & 0 \\ 0 & 0 \end{pmatrix}$ where $E_{(1)}$ is an $m \times m$ invertible diagonal matrix with $m < d$. Next, write $P = \begin{pmatrix} P_{(1)} & P_{(2)} \\ P_{(3)} & P_{(4)} \end{pmatrix}$ where $P_{(1)}$ is an $m \times m$ matrix. Since $PE = E$, we have

$$\begin{pmatrix} P_{(1)} E_{(1)} & 0 \\ P_{(3)} E_{(1)} & 0 \end{pmatrix} = \begin{pmatrix} E_{(1)} & 0 \\ 0 & 0 \end{pmatrix}.$$

From $P_{(1)} E_{(1)} = E_{(1)}$ and the invertibility of $E_{(1)}$, we have that $P_{(1)} = I_m$. Similarly, from $P_{(3)} E_{(1)} = 0$ we have that $P_{(3)} = 0$. Therefore, $P = \begin{pmatrix} I_m & P_{(2)} \\ 0 & P_{(4)} \end{pmatrix}$. Next, multiplying PP^t, and $P^t P$, and equating the $(1, 1)$ entries we have that $I_m + P_{(2)} P_{(2)}^t = I_m$. From this we conclude that $P_{(2)} P_{(2)}^t = 0$, and therefore also, $P_{(2)} = 0$. We have that,

$$P = \begin{pmatrix} I & 0 \\ 0 & P_{(4)} \end{pmatrix}.$$

The proof continues inductively. Let $K_{(n4)}, P_{(n4)}, E_{(n4)}$ be the (2,2) block of K_n, P_n, E_n respectively. $P_{(n4)}$ may not be orthogonal, but $P_{(4)}$ is. Apply the previous argument to $\left(K_{(n4)} P_{(n4)} P_{(4)}^t\right) P_{(4)} E_{(n4)}$. Note that $K_{(n4)} P_{(n4)} P_{(4)}^t \to I P_{(4)} P_{(4)}^t = I$, so that we may apply the argument with $K_{(n4)} P_{(n4)} P_{(4)}^t$ as the new K_n in the induction step. Since the matrices are all finite dimensional, the argument will eventually terminate.

Acknowledgement We would like to thank the referee for suggestions on how to shorten this article. The suggestions led to a much more efficient presentation of the material.

REFERENCES

[1] Billingsley, P. (1966). Convergence of types in k-space. *Z. Wahrsch. Verw. Gebiete.* **5** 175-179.

[2] Hahn, M. G. and Klass, M. J. (1980). Matrix normalization of sums of random vectors in the domain of attraction of the multivariate normal. *Ann. Probab.* **8** 262-280.

[3] Sepanski, S.J. (1993). Asymptotic normality of multivariate t and Hotelling's T^2 statistics under infinite second moments via bootstrapping. *J. Mutlivariate Analysis,* **49** 41-54

[4] Sepanski, S. J. (1994). Necessary and sufficient conditions for the multivariate bootstrap of the mean. *Statistics and Probability Letters,* **19** No. 3, 205-216

[5] Sepanski, S. J. (1996). Asymptotics for the multivariate t-statistic for random vectors in the generalized domain of attraction of the multivariate law. *Statistics and Probability Letters,* **30** 179-188.