

# A note on necessary and sufficient conditions for proving that a random symmetric matrix converges to a given limit

Robert L. Strawderman

*Department of Biostatistics, University of Michigan, 1420 Washington Heights, Ann Arbor, MI 48109-2029, USA*

Received August 1993; revised December 1993

---

## Abstract

We demonstrate that if  $B_n^{k \times k}$  is a sequence of symmetric matrices that converges in probability to some fixed but unspecified nonsingular symmetric matrix  $B$  elementwise, then  $B = B_0$  for a specified matrix  $B_0$  if and only if both the trace and squared Euclidean norm of  $D_n D_n^T$  converge to  $k$ , where  $D_n = B_0^{-1} B_n$ . Examples are given to demonstrate how this result may be used to construct hypothesis tests for the equality of covariance matrices and for model misspecification.

*Keywords:* Convergence in probability; Eigenvalues; Euclidean norm; Random matrix; Trace

---

## 1. Introduction

In the course of proving results involving the limiting behavior of matrices, one often needs to demonstrate that two matrices are equivalent in some sense. For example, we might be interested in demonstrating that the estimated covariance matrix of some statistic, say  $\hat{\Sigma}_n$ , converges to a given matrix  $\Sigma$ , or that two matrix sequences  $\hat{\Sigma}_{1n}$  and  $\hat{\Sigma}_{2n}$  converge to the same limit. There are many other types of problems which, at least to some degree, fit these general descriptions as well.

The purpose of this paper is to demonstrate that, under commonly made assumptions, problems like those outlined above may be studied through the behavior of two matrix functions that depend directly on matrix eigenvalues. In particular, we focus on the behavior of the trace and Euclidean norm; fortunately, both also have simple representations in terms of matrix elements, allowing one to apply the results without having to calculate any eigenvalues. Our results are formally presented in Section 2; some examples as to how these results may be applied are given in Section 3.

## 2. Results

Let  $X$  be a  $k \times k$  matrix,  $y$  be a  $k \times 1$  vector, and  $I$  be the  $k \times k$  identity matrix. The matrix and vector norms that we will be concerned with are defined as follows:  $\|X\|_\infty = \max_{i=1 \dots k} \sum_{j=1}^k |x_{ij}|$ ,  $\|X\|_E^2 = \sum_{i=1}^k \sum_{j=1}^k x_{ij}^2$ , and  $\|y\|_\infty = \max_{i=1 \dots k} |y_i|$ .

The following result will also be useful (Stewart, 1973, p. 181): let  $X_{1n}$  and  $X_{2n}$  be two sequences of square matrices such that  $\lim_{n \rightarrow \infty} \|X_{in} - X_i\|_\infty = 0, i = 1, 2$ , where  $X_i$  denote their respective limits; then,

$$\lim_{n \rightarrow \infty} \|X_{1n}X_{2n} - X_1X_2\|_\infty = 0. \tag{1}$$

We begin with a simple lemma:

**Lemma 1.** *Let  $A_n$  be a random sequence of symmetric nonnegative definite  $k \times k$  matrices where  $k < \infty$ . If a positive definite symmetric  $k \times k$  matrix  $A$  with finite elements exists such that  $A_n \xrightarrow{P} A$  elementwise, then  $\|A_n - A\| \xrightarrow{P} 0$ , where  $\|\cdot\|$  denotes any proper norm on  $\mathbb{R}^{k \times k}$ .*

**Proof.** By applying a sequence of elementary convergence results, it is easy to show that  $A_n \xrightarrow{P} A$  elementwise implies that  $\|A_n - A\|_E \xrightarrow{P} 0$ . Since all norms on  $\mathbb{R}^{k \times k}$  are equivalent (Golub and Van Loan, 1989, p. 58) and  $k < \infty$ , the desired result follows immediately.  $\square$

The converse of the above lemma is obviously true. We now provide the theorem and a useful corollary:

**Theorem 1.** *Let  $A_n$  be a random sequence of symmetric nonnegative definite  $k \times k$  matrices, where  $k < \infty$ , and assume that a positive definite symmetric  $k \times k$  matrix  $A$  exists such that  $A_n \xrightarrow{P} A$  elementwise and  $0 < \|A\|_\infty \leq M < \infty$ . Let  $\lambda_{jn}, j = 1, \dots, k$  denote the eigenvalues of  $A_n$ .*

*Then,  $A = I$  if and only if the following conditions both hold:*

- (i)  $\|A_n\|_E^2 = \sum_{j=1}^k \lambda_{jn}^2 \xrightarrow{P} k$
- (ii)  $\text{tr } A_n = \sum_{j=1}^k \lambda_{jn} \xrightarrow{P} k$

**Proof.** Let  $A_n = \Gamma_n A_n \Gamma_n'$ , where  $\Gamma_n = [\Gamma_{1n} \Gamma_{2n} \dots \Gamma_{kn}]$  is a  $k \times k$  orthonormal matrix of eigenvectors of  $A_n$ , and  $A_n = \text{diag}(\lambda_{1n}, \dots, \lambda_{kn})$  is the diagonal matrix of eigenvalues of  $A_n$ . Since  $A_n$  is assumed to be symmetric and nonnegative definite, such quantities exist and are well defined (Rao, 1965, pp. 36-37).

Let  $A = I$ . Then, given that  $A_n \xrightarrow{P} I$  elementwise, we know by Lemma 1 that  $\|A_n - I\|_E \xrightarrow{P} 0$ . Thus,  $\|A_n - I\|_E^2 \xrightarrow{P} 0$  by continuity. However, this implies that  $\sum_{j=1}^k (\lambda_{jn} - 1)^2 \xrightarrow{P} 0$  since  $\sum_{j=1}^k (\lambda_{jn} - 1)^2 \leq \|A_n - I\|_E^2$  for any square matrix  $A_n$  (Rao, 1965, p. 54). Thus, since  $\lambda_{jn} \in \mathbb{R}$  for  $j = 1, \dots, k, \lambda_{jn} \xrightarrow{P} 1, j = 1, \dots, k$ ; conditions (i) and (ii) follow by continuity.

To prove the theorem in the opposite direction, assume that (i) and (ii) hold. Then, by definition,  $\sum_{j=1}^k \lambda_{jn}^2 \xrightarrow{P} k$  and  $\sum_{j=1}^k \lambda_{jn} \xrightarrow{P} k$ . If we define  $\lambda_{jn} = 1 + \epsilon_{jn}$  for  $j = 1, \dots, k$ , where  $\lambda_{jn}, \epsilon_{jn} \in \mathbb{R}$ , then it is simple to show that  $\epsilon_{jn} \xrightarrow{P} 0$  and therefore that  $\lambda_{jn} \xrightarrow{P} 1$  for  $j = 1, \dots, k$ . Since  $\Gamma_n$  is an orthonormal matrix for any  $n$  (and hence  $\|\Gamma_{jn}\|_\infty \leq 1$  for any  $n, j = 1, \dots, k$ ), the spectral decomposition of  $A_n$ , given by  $A_n = \sum_{j=1}^k \lambda_{jn} \Gamma_{jn} \Gamma_{jn}'$ , must be asymptotically equivalent to the matrix  $B_n = \sum_{j=1}^k \Gamma_{jn} \Gamma_{jn}'$ . However, since  $B_n = I$  for any  $n$ , it follows that

$$(A_n)_{ij} \xrightarrow{P} \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases}$$

or  $A = I$ .  $\square$

A useful corollary to this theorem is as follows.

**Corollary 1.** *Let  $B_n$  be a sequence of symmetric  $k \times k$  matrices, and suppose that a matrix  $B$  exists such that  $B_n \xrightarrow{P} B$  elementwise where  $0 < \|B\|_\infty \leq M_1 < \infty$ . In addition, let  $B_0$  be a nonsingular symmetric  $k \times k$  matrix such that  $0 < \|B_0\|_\infty \leq M_2 < \infty$ . Then,  $B = B_0$  if and only if  $\|D_n D_n'\|_E^2 \xrightarrow{P} k$  and  $\text{tr } D_n D_n' \xrightarrow{P} k$ , where  $D_n = B_0^{-1} B_n$ .*

**Proof.** By Lemma 1, we know that  $B_n \xrightarrow{P} B$  is equivalent to  $\|B_n - B\|_\infty \xrightarrow{P} 0$ .

Now, let  $B = B_0$ . Using properties of norms, some straightforward algebra yields that

$$\|B_n^2 - B_0^2\|_\infty \|B_0^{-1}\|_\infty^2 \geq \|D_n D_n' - I\|_\infty.$$

By (1),  $\|B_n^2 - B_0^2\|_\infty \xrightarrow{P} 0$  since  $\|B_n - B_0\|_\infty \xrightarrow{P} 0$  by assumption. Thus, the left-hand side of the above inequality converges in probability to zero, and therefore so must the right-hand side since it is bounded below by zero. Hence, since  $D_n D_n'$  is symmetric and has limit  $I$ , the results of Theorem 1 imply that  $\|D_n D_n'\|_E^2 \xrightarrow{P} k$  and  $\text{tr } D_n D_n' \xrightarrow{P} k$ .

To prove the theorem in the opposite direction, assume that  $\|D_n D_n'\|_E^2 \xrightarrow{P} k$  and  $\text{tr } D_n D_n' \xrightarrow{P} k$ . By (1),  $\|D_n - D\|_\infty \xrightarrow{P} 0$  for  $D = B_0^{-1} B$  since  $\|B_n - B\|_\infty \xrightarrow{P} 0$ ; therefore,  $\|D_n D_n' - DD'\|_\infty \xrightarrow{P} 0$  by (1) as well. Hence,  $D_n D_n'$ , which is necessarily symmetric, satisfies the assumptions of Theorem 1. However, since  $\|D_n D_n'\|_E^2 \xrightarrow{P} k$  and  $\text{tr } D_n D_n' \xrightarrow{P} k$ , the results of Lemma 1 and Theorem 1 imply that  $\|D_n D_n' - I\|_\infty \xrightarrow{P} 0$  or that  $DD' = I$ . But this implies that  $B_0^{-1} B B B_0^{-1} = I$ , or that  $B = B_0$ , which is the desired result.  $\square$

### 3. Some examples

One application of these results is in constructing a test statistic for the equality of covariance matrices. For example, consider testing  $H_0: \Sigma_n = \Sigma_0$  using  $S_n^{k \times k}$  is an estimator of the population covariance matrix  $\Sigma_0$ . If we know the limiting distribution of  $n^{1/2}(S_n - \Sigma_0)$  under  $H_0$ , then we can determine the distribution of  $n^{1/2}(\Sigma_0^{-1} S_n - I)$ . Using the results of Eaton and Tyler (1991), we can thus determine the distribution of  $n^{1/2}(\psi(\Sigma_0^{-1} S_n) - \psi(I))$  where  $\psi(M)$  represents the vector of eigenvalues of a square matrix  $M$ . At this point, one often uses a test statistic based on the determinant ratio or perhaps the trace in order to test the null hypothesis. Given the results of the previous section, a potentially attractive alternative might be a quadratic form like

$$n(y - k\mathbf{1})' \hat{Q}^{2 \times 2} (y - k\mathbf{1}),$$

where

$$y = \sum_{j=1}^k \begin{pmatrix} \psi_j^4(\Sigma_0^{-1} S_n) \\ \psi_j^2(\Sigma_0^{-1} S_n) \end{pmatrix}$$

and  $\hat{Q}$  is an estimate of the asymptotic variance of  $n^{1/2}(y - k\mathbf{1})$ . Note that the above quadratic form is based on the symmetric matrix  $\Sigma_0^{-1} S_n S_n \Sigma_0^{-1}$ . Clearly, there are many variants to this statistic which may perform better than the one given. For example, using the trace and Euclidean norm of the symmetric matrix  $\Sigma_0^{-1/2} S_n \Sigma_0^{-1/2}$  may behave better in smaller samples, and using  $\log y$  instead of  $y$  may stabilize the variance. It would be interesting to compare the behavior of such a test to the tests currently being employed. Such a test may be easily extended to the two-sample problem of evaluating  $H_0: \Sigma_1 = \Sigma_2$ .

A related application involves constructing bootstrap tests for the equality of covariance matrices. For example, consider the two sample problem where we have two sets of vectors having equal dimension  $k$ , say  $x_1, \dots, x_n$  and  $y_1, \dots, y_m$ . Let us assume that they come from the same family of distributions, and we want to test  $H_0: \Sigma_x = \Sigma_y$ . One possible approach is obtain  $B$  bootstrap samples of the  $x$ 's and  $y$ 's, and for each resampled dataset calculate  $S_x^*$  and  $S_y^*$  and the trace and Euclidean norm of  $(S_y^*)^{-1} S_x^*$  (note: technically, one would need to look the trace and norm of a perturbation of  $(S_y^*)^{-1} S_x^*$  to account for the fact that there is one eigenvalue of multiplicity  $k$  under the null hypothesis; see, for example, Eaton and Tyler (1991) or Beran and Srivastava (1985)). Then, use the estimated joint distribution of the trace and Euclidean norm to see if the point  $(k, k)$  lies in the  $100(1 - \alpha)\%$ -level confidence set. If not, then reject the null hypothesis that  $\Sigma_x = \Sigma_y$ .

As a third example, consider the situation of testing model adequacy. White (1982) proposed a test based on the difference of the two obvious estimators of the Fisher information matrix, say  $\hat{\mathcal{J}}_1$  and  $\hat{\mathcal{J}}_2$ ; however, the resulting test statistic is compared to its asymptotic null distribution, and tends to perform poorly in general, even for moderate to large sample sizes. Thus, one possible approach is to look at the finite-sample behavior of the trace and Euclidean norm of  $\hat{\mathcal{J}}_2^{-1} \hat{\mathcal{J}}_1 \hat{\mathcal{J}}_2^{-1}$  using the parametric or nonparametric bootstrap. Then, as a test for model adequacy, we can use the estimated joint distribution of the trace and Euclidean norm to determine whether the point  $(k, k)$  lies in the  $100(1 - \alpha)\%$ -level confidence set. If not, then this provides evidence against the null hypothesis; if so, there is evidence to support the conclusion that the model has been adequately specified. Strawderman (1993) provides the details for the case of the censored data and uses the nonparametric bootstrap distribution of his test statistic to test for model misspecification; his methods directly apply to the uncensored case as well.

## References

- Beran, R. and M. Srivastava (1985), Bootstrap tests and confidence regions for functions of a covariance matrix, *Ann. Statist.* **13**, 95–113.
- Eaton, M.L. and D.E. Tyler (1991), On Wielandt's inequality and its application to the asymptotic distribution of the eigenvalues of a random symmetric matrix, *Ann. Statist.* **19**, 260–271.
- Golub, G. and C.F. Van Loan (1989), *Matrix Computations* (Johns Hopkins University Press, Baltimore, MD).
- Rao, C.R. (1965), *Linear Statistical Inference and Its Applications* (Wiley, New York, 1st ed).
- Serfling, R.J. (1980), *Approximation Theorems of Mathematical Statistics* (Wiley, New York).
- Stewart, G.W. (1973), *Introduction to Matrix Computations* (Academic Press, New York).
- Strawderman, R.L. (1993), Diagnosing model misspecification using the bootstrap: applications to censored data, unpublished manuscript.
- White, H. (1982), Maximum likelihood estimation of misspecified models, *Econometrica* **50**, 1–25.