Note on the convergence to normality of quadratic forms in independent variables

EINDHOVEN UNIVERSITY OF TECHNOLOGY
Department of Mathematics and Computing Science

COSOR memorandum 87-08

NOTE ON THE CONVERGENCE TO NORMALITY OF QUADRATIC FORMS IN INDEPENDENT VARIABLES

by

J.M. ten Vregelaar

Eindhoven University of Technology,
Department of Mathematics and Computing Science,
PO Box 513,
5600 MB Eindhoven,
The Netherlands.

Eindhoven, April 1987
The Netherlands

NOTE ON THE CONVERGENCE TO NORMALITY OF QUADRATIC FORMS IN INDEPENDENT VARIABLES

by
J.M. ten Vregelaar

ABSTRACT

In this paper we prove a result on the convergence to normality of quadratic forms in independent zero mean variables, which is a generalization of a theorem Whittle proved in his '64 paper (Whittle '64).

1. Introduction

In this paper we want to generalize a result on the convergence to normality of quadratic forms in independent variables (cf. Whittle '64).

The problem of determining limiting distributions of quadratic forms in random variables arises when considering asymptotic properties of least squares estimators for the parameters of an ARMA-model with noisy measurements of inputs and outputs (Ten Vregelaar '87).

Let us start giving a somewhat simplified version of Theorem 1 in Whittle '64.

Consider the quadratic form

$$x^T A x = \sum_{i,j=1}^{N^2} a_{ij} x_i x_j \tag{1.1a}$$

where

$$x_1, x_2, \cdots, x_{N^2} \text{ are independent random variables} \tag{1.1b}$$

and A is a $N^2 \times N^2$-symmetric matrix with elements a_{ij} depending on N.

We split up A as

$$A = \operatorname{diag}(A_1, A_2, \cdots, A_N) + A' , \tag{1.2a}$$

where A_i are $N \times N$ blocks, $i = 1,2, \cdots, N$.

The splitting is supposed to be "disjunct", in the sense that

$$\| A \|^2 = \sum_{i=1}^{N} \| A_i \|^2 + \| A' \|^2 , \tag{1.2b}$$

($\| \cdot \|$ denoting the Euclidean matrix norm).

If

$$\mathbb{E}\, x_i = 0 \quad \text{and} \quad \operatorname{var} x_i = \sigma^2, \qquad i = 1,2, \cdots, N^2 \tag{1.3}$$

the following conditions will be sufficient for $x^T A x$ to tend to normality in distribution with increasing N:

(ia) $\inf_i \mathbb{E}\, x_i^4 > \sigma^4$,

(ib) $\sup_i \mathbb{E}\, | x_i |^{4+\delta} < \infty$ for some $\delta > 0$,

(ii) $\sum_{i=1}^{N} \left[\frac{\| A_i \|}{\| A \|} \right]^{2+\delta} \xrightarrow{N \to \infty} 0$ for some $\delta > 0$

(iii) $\frac{\|A'\|}{\|A\|} \overset{N\to\infty}{\longrightarrow} 0.$

Condition (iii) expresses some diagonal dominance property. For sufficiently large N the distribution of the quadratic form $x^T A x$ approximates the distribution of

$$x^T \operatorname{diag}(A_1, A_2, \cdots, A_N) x \ ,$$

which can be written as a sum of independent random variables. Applying some central limit theorem, Condition (ii) guarantees, roughly spoken, the asymptotic normality of this sum.

The generalization we will consider is dropping the "disjunct" assumption (1.2b).

The proof goes along the same lines as Whittle's.

2. Result

Consider the quadratic form (1.1). Again we split up

$$A = \operatorname{diag}(A_1, A_2, \cdots, A_N) + A' \ , \tag{2.1}$$

now only assuming both A and A' (hence $A_1, \cdots, A_N$) are symmetric and $A_1, \cdots, A_N$ are $N \times N$ blocks.

Theorem.

The random variables $x_1, \cdots, x_{N^2}$ are independent and

$$\mathbb{E}\, x_i = 0\,, \ \ \operatorname{var} x_i = \sigma^2 \qquad (i = 1, 2, \cdots, N^2)\,. \tag{2.2}$$

The following conditions hold:

(ia) $\inf_i \mathbb{E}\, x_i^4 > \sigma^4$,

(ib) $\sup_i \mathbb{E}\, |x_i|^{4+\delta} < \infty$ for some $\delta > 0$,

the matrices $A_1, A_2, \cdots, A_N$ and A' in (2.1) can be chosen in such a way that

(ii) $\sum_{i=1}^{N} \left[\frac{\|A_i\|}{\|A\|} \right]^{2+\delta} \overset{N\to\infty}{\longrightarrow} 0$ for some $\delta > 0$,

and

(iii) $\frac{\|A'\|}{\|A\|} \overset{N\to\infty}{\longrightarrow} 0\,.$

Then the quadratic form $x^T A x$ is asymptotically normally distributed, i.e.

$$\frac{x^T A x - \mathbb{E}\, x^T A x}{\sqrt{\operatorname{var} x^T A x}}$$

converges in distribution to the standard normal distribution.

Remark.
For the special case $A_1 = A_2 = \cdots = A_N = \tilde{A}$, Condition (ii) follows from Condition (iii).

This can be shown by first proving a

Lemma.
Let P, Q and R be $N \times N$-matrices with

$$P = Q + R \ .$$

If

$$\frac{\| R \|}{\| P \|} \stackrel{N \to \infty}{\longrightarrow} 0 \ ,$$

then

$$\frac{\| Q \|}{\| P \|} \stackrel{N \to \infty}{\longrightarrow} 1 \ .$$

Proof. Since $Q = P - R$ we have

$$\| Q \|^2 = \| P \|^2 + \| R \|^2 - 2 \sum_{i,j=1}^{N} p_{ij} \, r_{ij} \ .$$

Applying the Cauchy-Schwartz inequality gives

$$\left| \sum_{i,j=1}^{N} p_{ij} \, r_{ij} \right| \leq \| P \| \| R \| \ .$$

The lemma follows from

$$\frac{\| Q \|^2}{\| P \|^2} = 1 + \frac{\| R \|^2}{\| P \|^2} - 2 \, \frac{\sum_{i,j=1}^{N} p_{ij} \, r_{ij}}{\| P \|^2}$$

since the modulus of the last term in the r.h.s. does not exceed $\frac{\| R \|}{\| P \|}$. ▯

Therefore Condition (iii) implies

$$\frac{\sqrt{N} \, \| \tilde{A} \|}{\| A \|} \stackrel{N \to \infty}{\longrightarrow} 1 \ ,$$

so

$$\sum_{i=1}^{N} \left[\frac{\| \tilde{A} \|}{\| A \|} \right]^{2+\delta} = N \left[\frac{\| \tilde{A} \|}{\| A \|} \right]^{2+\delta} = \left[\frac{\sqrt{N} \, \| \tilde{A} \|}{\| A \|} \right]^{2+\delta} \frac{1}{N^{\delta/2}} \stackrel{N \to \infty}{\longrightarrow} 0$$

for any $\delta > 0$.

3. Proof of the theorem

The proof is given by means of three lemmas.

Lemma A.

Let S_N, Σ_N and σ_N be random variables with

$$S_N = \Sigma_N + \sigma_N \tag{3.1a}$$

and

$$\frac{\operatorname{var}\sigma_N}{\operatorname{var}\Sigma_N} \overset{N\to\infty}{\longrightarrow} 0 \tag{3.1b}$$

then Σ_N is asymptotically normal implies S_N is asymptotically normal.

Proof. For the special case $\mathbb{E}\,\sigma_N = 0$ the proof can be found in Bernstein '26.

Define

$$S'_N := S_N - \mathbb{E}\,\sigma_N = \Sigma_N + \sigma'_N$$

with $\sigma'_N := \sigma_N - \mathbb{E}\,\sigma_N$.

Obviously

$$\mathbb{E}\,\sigma'_N = 0 \quad \text{and} \quad \operatorname{var}\sigma'_N = \operatorname{var}\sigma_N \ ,$$

hence

$$\frac{\operatorname{var}\sigma'_N}{\operatorname{var}\Sigma_N} \overset{N\to\infty}{\longrightarrow} 0 \qquad \text{from 3.1b .}$$

Therefore S'_N is asymptotically normal and thus S_N is, since

$$\frac{S_N - \mathbb{E}\,S_N}{\sqrt{\operatorname{var}S_N}} = \frac{S'_N - \mathbb{E}\,S'_N}{\sqrt{\operatorname{var}S'_N}} \ .$$

□

Remarks.

1. It is easy to verify that for $S_N = \Sigma_N + \sigma_N$ we have $\frac{\operatorname{var}\sigma_N}{\operatorname{var}\Sigma_N} \overset{N\to\infty}{\longrightarrow} 0$ if and only if $\frac{\operatorname{var}\sigma_N}{\operatorname{var}S_N} \overset{N\to\infty}{\longrightarrow} 0$, by applying the Cauchy-Schwartz inequality for random variables.
2. Unlike Whittle's proof, $\mathbb{E}\,\sigma_N = 0$ does not necessarily hold now.

We associate Lemma A and the theorem by defining

$$S_N := x^T A\, x\ , \ \ \sigma_N := x^T A' x\ , \ \ \Sigma_N := S_N - \sigma_N\ . \tag{3.2}$$

Lemma B.
Introduce

$$d := \frac{1}{\sigma^4} \inf_i \mathbb{E}\, x_i^4 .$$

If

1. $d > 0$,
2. $\sup_i \mathbb{E}\, x_i^4 < \infty$,
3. $\frac{\|A'\|}{\|A\|} \stackrel{N\to\infty}{\longrightarrow} 0$

then

$$\frac{\operatorname{var}\sigma_N}{\operatorname{var} S_N} \stackrel{N\to\infty}{\longrightarrow} 0 .$$

Proof. From Relation (18) in Whittle '64 (p. 106) we obtain

$$\operatorname{var} S_N \geq \min(2,d)\, \sigma^4 \|A\|^2 , \tag{3.3}$$

and Theorem 2 in Whittle '60 (p. 302) implies

$$\operatorname{var}\sigma_N \leq K \sup_i \mathbb{E}\, x_i^4 \|A'\|^2 \quad \text{for some constant } K .$$

Consequently,

$$\frac{\operatorname{var}\sigma_N}{\operatorname{var} S_N} \leq \frac{K \sup_i \mathbb{E}\, x_i^4}{\sigma^4 \min(2,d)} \; \frac{\|A'\|^2}{\|A\|^2} \stackrel{N\to\infty}{\longrightarrow} 0$$

provided Conditions 1-3. □

Remark.
In case the vector x has a multinormal distribution, ($x \sim \mathbf{N}(0, \sigma^2 I)$)

$$\frac{\operatorname{var}\sigma_N}{\operatorname{var} S_N} = \frac{2\sigma^4\|A'\|^2}{2\sigma^4\|A\|^2} = \frac{\|A'\|^2}{\|A\|^2}$$

holds.

Let us consider now the quadratic form Σ_N. From (2.1) and (3.2) it is obvious that Σ_N can be written as

$$\Sigma_N = \sum_{i=1}^{N} Y_i , \qquad Y_1, Y_2, \cdots, Y_N \text{ independent} . \tag{3.4}$$

The Liapounov central limit theorem applies to obtain asymptotic normality for Σ_N (cf. Serfling '80, p. 30):

Lemma C.

If for some $\delta > 0$, the conditions

1. $\mathbb{E}\ |Y_i - \mathbb{E}\ Y_i|^{2+\delta} < \infty$ for $i = 1,2,\cdots,N$,

2. $\sum_{i=1}^{N} \mathbb{E} \left[\frac{|Y_i - \mathbb{E}\ Y_i|}{\sqrt{\sum_{i=1}^{N} \operatorname{var} Y_i}} \right]^{2+\delta} \overset{N\to\infty}{\longrightarrow} 0$

are satisfied, then $\frac{\Sigma_N - \mathbb{E}\ \Sigma_N}{\sqrt{\operatorname{var}\Sigma_N}}$ converges in distribution to the standard normal distribution.

Now we are able to prove the theorem in Section 2. Since $\Sigma_N = S_N - \sigma_N$,

$$\operatorname{var}\Sigma_N = \operatorname{var} S_N + \operatorname{var}\sigma_N - 2\operatorname{cov}(S_N,\sigma_N) =$$

$$= \operatorname{var} S_N \left[1 + \frac{\operatorname{var}\sigma_N}{\operatorname{var} S_N} - 2\frac{\operatorname{cov}(S_N,\sigma_N)}{\operatorname{var} S_N} \right]$$

holds, where

$$\frac{|\operatorname{cov}(S_N,\sigma_N)|}{\operatorname{var} S_N} \leq \left[\frac{\operatorname{var}\sigma_N}{\operatorname{var} S_N} \right]^{\frac{1}{2}} \tag{3.5}$$

(from Cauchy-Schwartz).

Then

$$\sum_{i=1}^{N} \mathbb{E} \left[\frac{|Y_i - \mathbb{E}\ Y_i|}{\sqrt{\operatorname{var}\Sigma_N}} \right]^{2+\delta} = \sum_{i=1}^{N} \mathbb{E} \left[\frac{|Y_i - \mathbb{E}\ Y_i|}{\sqrt{\operatorname{var} S_N}\sqrt{1 + \frac{\operatorname{var}\sigma_N}{\operatorname{var} S_N} - 2\frac{\operatorname{cov}(S_N,\sigma_N)}{\operatorname{var} S_N}}} \right]^{2+\delta} .$$

Hence

$$\frac{\sum_{i=1}^{N} \mathbb{E} \left[\frac{|Y_i - \mathbb{E}\ Y_i|}{\sqrt{\operatorname{var}\Sigma_N}} \right]^{2+\delta}}{\sum_{i=1}^{N} \mathbb{E} \left[\frac{|Y_i - \mathbb{E}\ Y_i|}{\sqrt{\operatorname{var} S_N}} \right]^{2+\delta}} = \frac{1}{\left[1 + \frac{\operatorname{var}\sigma_N}{\operatorname{var} S_N} - 2\frac{\operatorname{cov}(S_N,\sigma_N)}{\operatorname{var} S_N} \right]^{1+\delta/2}} . \tag{3.6}$$

From Lemma B, the Conditions (i) and (iii) of the theorem are sufficient to obtain

$$\operatorname{var}\frac{\sigma_N}{S_N} \overset{N\to\infty}{\longrightarrow} 0 , \tag{3.7}$$

so the l.h.s. of (3.6) tends to 1 for $N \to \infty$ (using (3.5)).

Let us now investigate the nominator of this expression. Notice that we can write Y_i in (3.4) as

$$Y_i = \bar{x}_i^T A_i \bar{x}_i \ ,$$

with

$$x = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_N \end{bmatrix}, \qquad \bar{x}_i \text{ columnvector of } N \text{ components .}$$

Applying Theorem 2 in Whittle '60, p. 302, again ($s = 2+\delta$), yields

$$\mathbb{E}\ |Y_i - \mathbb{E}\ Y_i|^{2+\delta} \le C \sup_j \mathbb{E}\ |\bar{x}_{i_j}|^{4+2\delta} \|A_i\|^{2+\delta} \le$$

$$\le C \sup_i \mathbb{E}\ |x_i|^{4+2\delta} \|A_i\|^{2+\delta} < \infty \qquad \text{for all } i$$

(C is some constant), so Condition 1 in Lemma C is satisfied.

Combining this with (3.3) gives for the nominator of the l.h.s. in (3.6):

$$\sum_{i=1}^{N} \mathbb{E} \left[\frac{|Y_i - \mathbb{E}\ Y_i|}{\sqrt{\operatorname{var} S_N}} \right]^{2+\delta} = \frac{\sum_{i=1}^{N} \mathbb{E}\ |Y_i - \mathbb{E}\ Y_i|^{2+\delta}}{(\operatorname{var} S_N)^{1+\delta/2}} \le$$

$$\le \frac{C \sup_i \mathbb{E}\ |x_i|^{4+2\delta}}{(\min(2,d))^{1+\delta/2} \sigma^{4+2\delta}} \sum_{i=1}^{N} \left[\frac{\|A_i\|}{\|A\|} \right]^{2+\delta} .$$

Provided Conditions (i), (ii) of the theorem, the nominator of the l.h.s. of (3.6) tends to 0 for $N \to \infty$.

Hence the denominator of this term, being the product of the nominator and the l.h.s. of (3.6) itself, converges to 0 for $N \to \infty$. The conditions in Lemma C are satisfied, we have proved so far the asymptotic normality of Σ_N.

The theorem is proved now by applying Lemma A, since (3.1b) is a consequence of (3.7), according to Remark 1 at Lemma A.

4. Discussion

The theorem will remain valid, if we replace (2.2) by

$$\mathbb{E}\ x_i = 0\,, \quad \operatorname{var} x_i = \sigma_i^2$$

where

$$0 < a \le \sigma_i^2 \le b < \infty\ , \qquad i = 1,2, \cdots, N^2\ .$$

We can even drop the latter by introducing the norm as Whittle did (Whittle '64, p. 105).

Furthermore, it is possible to split up the set of integers more generally as

$$\{1,2, \cdots ,N^2\} = \bigcup_{i=1}^{r} g_i$$

where g_i are distinct sets consisting of a varying number of integers (cf. Whittle '64).

Condition (ia) only excludes some pathological distributions like

$$\mathbb{P}(x_j = 1) = \tfrac{1}{2}$$

$$\mathbb{P}(x_j = -1) = \tfrac{1}{2}$$

which satisfies $\mathbb{E}\, x_i^2 = \mathbb{E}\, x_i^4 = 1$.

References

Bernstein '26 — Bernstein, S.N., Sur l'extension du théorème limite du calcul des probabilités aux sommes de quantités dépendantes, Math. Ann., 97, pp. 1-59, 1926.

Serfling '80 — Serfling R.J., Approximation Theorems of Mathematical Statistics, John Wiley and Sons, New York, 1980.

Ten Vregelaar '87 — Ten Vregelaar, J.M., Asymptotic properties of some estimation method for an ARMA-model with noisy measurements of inputs and outputs, Memorandum-COSOR, Eindhoven University of Technology, in preparation.

Whittle '60 — Whittle, P., Bounds for the moments of linear and quadratic forms in independent variables, Theory Prob. Applications, 5, pp. 302-305, 1960.

Whittle '64 — Whittle, P., On the convergence to normality of quadratic forms in independent variables, Theory Prob. Applications, 9, 1, pp. 103-108, 1964.