ASYMPTOTIC DISTRIBUTION OF QUADRATIC FORMS II

F. Götze[1] and A. Tikhomirov[1,2]

University of Bielefeld[1]
Syktyvkar State University and the Mathematical Department
of Komi RC of the Russian Academy of Sciences[2]

November 1996

Abstract. We consider the quadratic forms

$$Q_n = \sum_{1 \le j,k \le n} a_{jk}\big(X_j X_k - \mathbf{E}\, X_j X_k\big),$$

where X_j are i.i.d. random variables with finite sixth moment. For a large class of matrices a_{jk} the distribution of Q_n can be approximated by the distribution of a second order polynomial in Gaussian random variables. We provide optimal bounds for the Kolmogorov distance of these distributions, extending previous results for matrices with zero diagonals to the general case.

1. Introduction and results

Let $A = A^{(n)} = \{a_{jk}^{(n)}\}_{j,k=1}^{n}$ denote a sequence of symmetric matrices such that

$$\sum_{j,k=1}^{n} \big(a_{jk}^{(n)}\big)^2 = 1. \tag{1.1}$$

In order to simplify the notation we will drop in the following, the index n.

$X, X_1, X_2, \dots$ denote independent identically distributed (i.i.d.) random variables (r.v.) such that $\mathbf{E}\, X = 0$. Consider the quadratic form

$$Q = \sum_{j,k=1}^{n} a_{jk}\big(X_j X_k - \mathbf{E}\, X_j X_k\big).$$

1991 *Mathematics Subject Classification.* 60F05.

Key words and phrases. independent random variables, quadratic forms, asymptotics of distribution, limit theorems, Berry–Esseen bounds.

[1]Research supported by the SFB 343 at Bielefeld.

[2]Partially supported by Russian Foundation for Fundamental Research. Grant N96–01–00201. Partially supported by INTAS-RFBR, DFG-RFBR.

The aim of this paper is to bound the error of approximation of the distribution of the quadratic form Q by that of a suitable polynomial of second order in Gaussian random variables, for sufficiently general classes of matrices.

Introduce $\mu_k = \mathbf{E}\, X^k$ and assume that

$$\mu_6 < \infty, \tag{1.2}$$

and $\mu_2 > 0$. Let $Y, Y_1, Y_2, \dots$ denote a sequence of independent Gaussian random variables. Consider the polynomial

$$G = \mu_2 \sum_{1\le j\ne k\le n} a_{jk} Y_j Y_k + \mu_3 \mu_2^{-1/2} \sum_{j=1}^{n} a_{jj} Y_j + \sqrt{\mu_4 - \mu_2^2 - \mu_3^2 \mu_2^{-1}}\, V Y,$$

where $V^2 = \sum\limits_{j=1}^{n} a_{jj}^2$. The inequality $|\mathbf{E}\,(X^2 - \mathbf{E}\,X^2)X|^2 \le \mathbf{E}\,(X^2 - \mathbf{E}\,X^2)^2 \mathbf{E}\,X^2$ implies that $\mu_4 - \mu_2^2 - \mu_3^2\mu_2^{-1} \ge 0$. Without loss of generality we may assume that $\mu_2 = 1$. Let us denote

$$\mathcal{L}_j^2 = \sum_{\substack{k\ne j\\ k=1,\dots,n}} a_{jk}^2, \qquad \mathcal{L}^2 = \max_{1\le j\le n} \mathcal{L}_j^2, \qquad \mathcal{M} = \max_{1\le j\le n} |a_{jj}|,$$

and

$$\delta(A,F) := \sup_x \Big| \mathbf{P}\,\big\{ Q \le x \big\} - \mathbf{P}\,\big\{ G \le x \big\} \Big|.$$

We will write $\Phi(x)$ for the standard Gaussian distribution function and denote

$$\Delta(A,F) := \sup_x \Big| \mathbf{P}\,\big\{ Q \le x \big\} - \Phi(x) \Big|,$$

where F is the distribution function (d.f.) of X, i.e. $F(x) = \mathbf{P}\,\{X \le x\}$. Throughout this paper C (with index or without it) denote absolute constants, whereas $C(\cdot,\cdot)$ is used for constants depending on the arguments in parentheses. Introduce the matrix A_0 with elements $A_0(j,k) = a_{jk}$ for $1 \le j \ne k \le n$, and $A_0(j,j) = 0$. By λ_k, $k = 1, 2, \dots, n$, we denote the eigenvalues of the matrix A_0 ordered so that $|\lambda_1| \ge \dots \ge |\lambda_n|$. Note that, by (1.1), $\sum\limits_{k=1}^{n} \lambda_k^2 = 1$.

Define $\log^+ x = \max\{1/2, |\log x|\}$, $x > 0$. The main results of this paper are contained in the following theorems.

Theorem 1. *Assume that the conditions* (1.1)–(1.2) *hold. Then there exists an absolute constant C such that*

$$\max\big\{\, \delta(A,F), \Delta(A,F) \,\big\} \le C\,\mu_6 \Big(\big|\lambda_1\big| + \mathcal{M} \Big).$$