PLISKA

STUDIA MATHEMATICA BULGARICA

ПЛИСКА

БЪЛГАРСКИ МАТЕМАТИЧЕСКИ СТУДИИ

For further information on
Pliska Studia Mathematica Bulgarica
visit the website of the journal http://www.math.bas.bg/~pliska/
or contact: Editorial Office
Pliska Studia Mathematica Bulgarica
Institute of Mathematics and Informatics
Bulgarian Academy of Sciences
Telephone: (+359-2)9792818, FAX:(+359-2)971-36-49
e-mail: pliska@math.bas.bg

Pliska Stud. Math. Bulgar. **22** (2013), 225–236

PLISKA
STUDIA MATHEMATICA
BULGARICA

MARGINAL DENSITIES OF THE WISHART DISTRIBUTION

Evelina Veleva

ABSTRACT. We consider marginal densities obtained by elimination of non-diagonal elements of a positive definite random matrix with an arbitrary distribution. For a $p \times p$ random matrix $\mathbf{W}$ such a marginal density is presented by a graph with p vertices. For every non-diagonal element of $\mathbf{W}$, included in the density we draw in the graph an undirected edge between the corresponding vertices. By giving an equivalent definition of decomposable graphs we show that the bounds of the integration with respect to every excluded element of $\mathbf{W}$ can be exactly obtained if and only if the corresponding graph is decomposable. The author gives in an explicit form some of the marginal densities of an arbitrary Wishart distribution.

1. Introduction. Wishart distribution arises as the distribution of the sample covariance matrix for a sample from a multivariate normal distribution. Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be n independent observations on a random vector $\mathbf{x}$ with p-variate normal distribution $N_p(\mu, \Sigma)$, $p < n$, with mean vector μ and positive definite covariance matrix Σ. Let $\mathbf{S}$ be the sample covariance matrix $\mathbf{S} = \left(\sum_{i=1}^{n} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^t\right)/(n-1)$, $\bar{\mathbf{x}} = \left(\sum_{i=1}^{n} \mathbf{x}_i\right)/n$. Then the joint distribution of the elements of the matrix $\mathbf{S}$ is Wishart distribution $\mathrm{W}_p\left(n-1,\ 1/(n-1)\Sigma\right)$ (see [1], [5]). Hence the joint distributions of sets of elements of the matrix $\mathbf{S}$ are marginal distributions of the Wishart distribution.

2010 *Mathematics Subject Classification:* 62H10.

Key words: Wishart distribution, positive definite matrix, marginal density, covariance matrix, decomposable graph.

Let $\mathbf{W} = (W_{i,j})$ be a random matrix with Wishart distribution $\mathrm{W}_p(n, \Sigma)$. Let $\mathbf{W}[\{k, \dots, s\}]$ denotes the submatrix of the matrix $\mathbf{W}$, composed of the rows and columns with numbers from the set $\{k, \dots, s\}$, where $1 \leq k \leq s \leq p$. It is known (see [1], [5]) that $\mathbf{W}[\{k, \dots, s\}]$ has again Wishart distribution $\mathrm{W}_{s-k+1}(n, \Sigma[\{k, \dots, s\}])$. This marginal distribution corresponds to all the elements $\{W_{i,j}, k \leq i \leq j \leq s\}$ on and above the main diagonal in the submatrix $\mathbf{W}[\{k, \dots, s\}]$. Marginal densities for the sets of the form

$$\{W_{i,j}, k \leq i \leq j \leq s\} \backslash \{W_{q,r}\}, \tag{1}$$

where $k \leq q < r \leq s$, are obtained in [6] by integration the density of the Wishart distribution (2) below with respect to the element $w_{q,r}$ of the positive definite matrix W. In this paper we give alternative representations of these marginal densities and obtain some generalizations. We consider marginal densities obtained by elimination of non-diagonal elements of a positive definite random matrix with an arbitrary distribution. For a $p \times p$ random matrix $\mathbf{W}$ such a marginal density is presented by a graph with p vertices. For every non-diagonal element of $\mathbf{W}$, included in the density we draw in the graph an undirected edge between the corresponding vertices. By giving an equivalent definition of decomposable graphs we show that the bounds of the integration with respect to every excluded element of $\mathbf{W}$ can be exactly obtained if and only if the corresponding graph is decomposable.

2. Preliminary notes. A $p \times p$ random matrix $\mathbf{W}$ with Wishart distribution $\mathrm{W}_p(n, \Sigma)$, where $p < n + 1$ and Σ is a positive definite $p \times p$ matrix, has probability density function of the form

$$f_{p,n,\Sigma}(\mathrm{W}) = \frac{1}{2^{np/2}\Gamma_p(n/2)(\det \Sigma)^{n/2}} (\det \mathrm{W})^{(n-p-1)/2} e^{-tr(\mathrm{W}\Sigma^{-1})/2} \tag{2}$$

for any real $p \times p$ positive definite matrix W, where $\Gamma_p(\cdot)$ is the multivariate gamma function defined as $\Gamma_p(\gamma) = \pi^{p(p-1)/4} \prod_{j=1}^{p} \Gamma[\gamma + (1-j)/2]$ and $\det(\cdot)$, $tr(\cdot)$ denote the determinant and the trace of a matrix.

For a given $p \times p$ matrix W, by $\mathrm{W}[\alpha, \beta]$ we denote the submatrix of W, composed of the rows with numbers from the set α and the columns with numbers from the set β, where α, β are nonempty subsets of the set $\{1, \dots, p\}$.

When $\beta \equiv \alpha$, $\mathrm{W}[\alpha, \alpha]$ is denoted simply by $\mathrm{W}[\alpha]$. For the complement of α in $\{1, \dots, p\}$ we use the notation α^c.

For instance, $\mathrm{W}[\{q\}^c, \{r\}^c]$ denotes the submatrix, which can be obtained from W by deleting its q-th row and r-th column.

The next Proposition, which is proved in [6], gives the bounds for the integration of the density function of a random positive definite matrix with an arbitrary distribution with respect to an arbitrary chosen its non – diagonal element.

Proposition 1. *Let* $\mathrm{W} = (w_{i,j})$ *be a real* $p \times p$ *symmetric matrix and* q, r *be fixed integers,* $1 \leq q < r \leq p$. *Let* W_0 *be the matrix, obtained from the matrix* W *by replacing the elements* $w_{q,r}$ *and* $w_{r,q}$ *with zeros. The matrix* W *is positive definite if and only if the matrices* $\mathrm{W}[\{q\}^c]$ *and* $\mathrm{W}[\{r\}^c]$ *are positive definite and the element* $w_{q,r}$ *satisfies the inequalities*

$$A - B < w_{q,r} < A + B$$

where

$$A = \frac{(-1)^{r-q} \det \mathrm{W}_0[\{q\}^c, \{r\}^c]}{\det \mathrm{W}[\{q, r\}^c]}, \quad B = \frac{\sqrt{\det \mathrm{W}[\{q\}^c] \det \mathrm{W}[\{r\}^c]}}{\det \mathrm{W}[\{q, r\}^c]}. \tag{3}$$

The result after the integration of the Wishart density (2) is given by the next Proposition (see [6]) in terms of modified Bessel function of the first kind $I_v(\cdot)$ (see [2], 8.445). Throughout the paper, the elements of the matrix Σ^{-1} are denoted by $\sigma^{i,j}$, $1 \leq i \leq j \leq p$.

Proposition 2. *Let* $\mathbf{W} = (W_{i,j})$ *has Wishart distribution* $\mathrm{W}_p(n, \Sigma)$ *and* q, r *be integers,* $1 \leq q < r \leq p$. *Then the marginal density, corresponding to the set of elements* $\{W_{i,j},\ 1 \leq i \leq j \leq p\} \backslash \{W_{q,r}\}$ *has the form*

$$f_{\{W_{q,r}\}^c}(\mathrm{W}_0) = \frac{L}{2^{np/2} \Gamma_p(n/2) (\det \Sigma)^{n/2}} \frac{(\det \mathrm{W}_0[\{q\}^c] \det \mathrm{W}_0[\{r\}^c])^{(n-p)/2}}{(\det \mathrm{W}_0[\{q, r\}^c])^{(n-p+1)/2}} e^{-tr(\mathrm{W}_0 \Sigma^{-1})/2}$$

for every symmetric $p \times p$ *matrix* $\mathrm{W}_0 = (w_{i,j})$, *such that* $w_{q,r} = w_{r,q} = 0$ *and the matrices* $\mathrm{W}_0[\{q\}^c]$, $\mathrm{W}_0[\{r\}^c]$ *are both positive definite. If* $\sigma^{q,r} = 0$, *then*

$$L = \frac{\Gamma\left((n-p+1)/2\right)\ \Gamma(1/2)}{\Gamma\left((n-p+2)/2\right)}. \tag{4}$$

For $\sigma^{q,r} \neq 0$,

$$L = \Gamma\left((n-p+1)/2\right)\ \Gamma(1/2)\ e^{-A\sigma^{q,r}} \left(\frac{2}{B\sigma^{q,r}}\right)^{(n-p)/2} I_{(n-p)/2}(B\sigma^{q,r}) \tag{5}$$

where A *and* B *are given with* (3).

Let $\mathbf{W} = (W_{i,j})$ be a $p \times p$ positive definite random matrix with an arbitrary distribution. Let M be a subset of the set $\{W_{i,j}, 1 \le i < j \le p\}$ of non-diagonal elements of $\mathbf{W}$. The marginal density f_{M^c}, obtained after integration of the density function of $\mathbf{W}$ with respect to the variables from M, can be presented by a graph $G_{\_M}$ with the set of vertices $V = \{1, \dots, p\}$. For every element $W_{i,j}$ of the set $M^c = \{W_{i,j}, 1 \le i < j \le p\} \backslash M$ we draw in the graph an undirected edge between the vertices "i" and "j". The same idea is used in [7] to describe joint densities of sample correlation coefficients.

We shall recall some definitions and properties of graphical models. These standard ideas are covered more thoroughly in [4].

Consider a graph $G = (V, E)$ with a finite set of vertices V and a set of undirected edges E. G is called a complete graph if every pair of distinct vertices is connected by an edge. A subset of vertices $U \subseteq V$ defines an induced subgraph G_U of G which contains all the vertices U and any edges in E that connect vertices in U. A clique is a complete subgraph that is maximal, that is, it is not a subgraph of any other complete subgraph.

Definition 1. *A graph G is decomposable if and only if the set of cliques of G can be ordered as $(C_1, \dots, C_k)$ so that for each $i = 2, \dots, k$ if $S_i = C_i \cap \bigcup_{j=1}^{i-1} C_j$ then $S_i \subset C_l$ for some $l < i$.*

Decomposable graphs are also known as triangulated or chordal graphs; Definition 1 is equivalent to the requirement that G contains no chordless cycles of length greater than 3. The next Proposition for decomposable graphs is given in [3].

Proposition 3. *Disconnecting x and y by removing an edge (x, y) from a decomposable graph G will result in a decomposable graph if and only if x and y are contained in exactly one clique.*

The next Proposition, proved in [8], is a generalization of the Sylvester's determinant identity.

Proposition 4. *Let A be a square matrix of order n and let i, k, j, l be integers, such that $1 \le i < k \le n$, $1 \le j < l \le n$. Then*

$$\det \mathrm{A} \det \mathrm{A}[\{i,k\}^c, \{j,l\}^c] = \det \mathrm{A}[\{i\}^c, \{j\}^c] \det \mathrm{A}[\{k\}^c, \{l\}^c] - \det \mathrm{A}[\{i\}^c, \{l\}^c] \det \mathrm{A}[\{k\}^c, \{j\}^c].$$

3. Main results. Let $\mathbf{W} = (W_{i,j})$ be a $p \times p$ random matrix with Wishart distribution $\mathrm{W}_p(n, \Sigma)$. The joint distribution of all $W_{i,j}$, $1 \le i \le j \le p$ corresponds to a complete graph $G = (V, E)$ with set of vertices $V = \{1, \ldots, p\}$ and the set of edges E, connecting every pair of distinct vertices in V. The integration of the density function (2) with respect to a variable $w_{q,r}$, where $1 \le q < r \le p$, corresponds to removing in the graph G the edge connecting the vertices "q" and "r". The resulting graph G_1 has two cliques C_1 and C_2, $C_1 = G_{V \setminus \{q\}}$ and $C_2 = G_{V \setminus \{r\}}$. According to Definition 1, G_1 is a decomposable graph. The obtained after the integration marginal density, given by Proposition 2, is defined for all elements of the matrix W_0, for which the submatrices corresponding to the two cliques C_1 and C_2, i.e. $\mathrm{W}_0[\{q\}^c]$ and $\mathrm{W}_0[\{r\}^c]$, are both positive definite.

Let the next integration be with respect to variable $w_{i,j}$, such that the edge (i, j) belongs to exactly one of the cliques C_1 and C_2, i.e. $(i, j) \notin C_1 \cap C_2$. These are variables for which one of the indices is q or r. Then $w_{i,j}$ is an element of exactly one of the matrices $\mathrm{W}_0[\{q\}^c]$ and $\mathrm{W}_0[\{r\}^c]$. Consequently, the bounds of the integration with respect to $w_{i,j}$ can be obtained by directly applying Proposition 1 to the matrix $\mathrm{W}_0[\{q\}^c]$ or $\mathrm{W}_0[\{r\}^c]$ where $w_{i,j}$ belongs.

If, however, the next integration is with respect to variable $w_{k,s}$ such that $\{q, r\} \cap \{k, s\} = \emptyset$, it will present in both the matrices $\mathrm{W}[\{q\}^c]$ and $\mathrm{W}[\{r\}^c]$. For any of the two matrices, Proposition 1 gives bounds for $w_{k,s}$ of the form

$$a_i < w_{k,s} < b_i, \quad i = 1, 2.$$

Then the integration with respect to $w_{k,s}$ have to be done in bounds from $\max(a_1, a_2)$ to $\min(b_1, b_2)$. The obtained after this integration marginal density will be defined for all $w_{i,j}$, $1 \le i \le j \le p$, $(i, j) \ne (q, r), (k, s)$ for which the matrices $\mathrm{W}[\{q, k\}^c]$, $\mathrm{W}[\{q, s\}^c]$, $\mathrm{W}[\{r, k\}^c]$ and $\mathrm{W}[\{r, s\}^c]$ are positive definite and

$$\max(a_1, a_2) < \min(b_1, b_2). \tag{6}$$

The inequality (6) can be solved only numerically with respect to a third variable which we would want to exclude.

Theorem 1. *Let $\mathbf{W} = (W_{i,j})$ be a $p \times p$ positive definite random matrix with an arbitrary distribution and M be a subset of the set $\{W_{i,j}, 1 \le i < j \le p\}$ of non-diagonal elements of $\mathbf{W}$. Then the bounds of the integration of the density function of $\mathbf{W}$ with respect to all the elements of M can be exactly obtained by directly applying Proposition 1 if and only if the corresponding graph $G_{\_M}$ is decomposable.*

To prove Theorem 1 we first give an equivalent definition for decomposable graphs.

Definition 2. *A graph G is decomposable if and only if it is a complete graph or it can be obtained from a complete graph by stepwise removing of edges such that each removed edge at the moment of its elimination belongs to exactly one clique.*

Lemma 1. *Definitions* 1 *and* 2 *for decomposable graphs are equivalent.*

Proof of Lemma 1. Let at first G be a decomposable graph according to Definition 1. We shall prove that G satisfies the condition in Definition 2. We shall use mathematical induction on the number k of the cliques of G. Let us assume at first that G has $k = 2$ cliques C_1 and C_2, and $(C_1 \cap C_2) \subset C_1$. Let us denote by V_i the set of vertices of the subgraph C_i, $i = 1, 2$. Let G' be the complete graph with set of vertices $V_1 \cup V_2$. Then G can be obtained from G' by removing the edges, connecting each vertex from the set $V_1 \backslash V_2$, i.e. $V_1 \cap \overline{V_2}$, with every vertex from the set $V_2 \backslash V_1$. The condition in Definition 2 will be satisfied if the elimination of the edges is done in the following sequence: we choose an arbitrary vertex v_1 from the set $V_1 \backslash V_2$ and one after the other we remove every edge, connecting v_1 with a vertex from the set $V_2 \backslash V_1$; then we repeat this procedure for another arbitrary chosen vertex from $V_1 \backslash V_2$ and so on until we do this for every vertex from the set $V_1 \backslash V_2$. After each removing of an edge connecting the first vertex v_1 with a vertex from the set $V_2 \backslash V_1$, the resulting graph always has two cliques one of which is $G'_{(V_1 \cup V_2) \backslash \{v_1\}}$, i.e. every time the next edge to remove belongs to the other clique. After the last removing of an edge connecting v_1 with a vertex from $V_2 \backslash V_1$, the resulting graph has again two cliques – one is $G'_{(V_1 \cup V_2) \backslash \{v_1\}}$ and the other is $C_1 = G'_{V_1}$. Then we repeat this procedure for another vertex v_2 from the set $V_1 \backslash V_2$. After each removing of an edge connecting v_2 with a vertex from $V_2 \backslash V_1$, the resulting graph always has three cliques two of which are $C_1 = G'_{V_1}$ and $G'_{(V_1 \cup V_2) \backslash \{v_1, v_2\}}$. Every time the next edge to remove belongs to the third clique. After the last removing of an edge connecting v_2 with a vertex from $V_2 \backslash V_1$, the resulting graph has two cliques – $C_1 = G'_{V_1}$ and $G'_{(V_1 \cup V_2) \backslash \{v_1, v_2\}}$. The removing of the edges connecting the rest vertices from the set $V_1 \backslash V_2$ with vertices from $V_2 \backslash V_1$ runs analogously; every time the removed edge belongs to exactly one clique. Finally, after the last removed edge we obtain the graph G with the two cliques $C_1 = G'_{V_1}$ and $C_2 = G'_{V_2}$. Let us assume now that every decomposable, according to Definition 1, graph with k cliques satisfies the condition in Definition 2. Let G be a decomposable, according to Definition 1, graph with $k+1$ cliques $C_1, \ldots, C_{k+1}$. We shall prove that G satisfies the condition in Definition 2. Let us denote by G_1 and G_2 the two subgraphs $G_1 = C_{k+1}$ and $G_2 = C_1 \cup \ldots \cup C_k$. According to Definition 1, G_2 is a decomposable graph and $G_1 \cap G_2$ is contained in some C_l, $l \leq k$. Let us denote by V_i the set of vertices of the subgraph G_i, $i = 1, 2$. Let G' be the complete graph with the set of vertices $V_1 \cup V_2$. Let us remove at first in G' the edges connecting each vertex from the set $V_1 \backslash V_2$ with every vertex from the set $V_2 \backslash V_1$. As it was already shown for the case $k = 2$, this can be done by stepwise removing of edges such that each removed edge at the moment of its elimination belongs to exactly one clique. The resulting graph will have two cliques - $G_1 = C_{k+1} = G'_{V_1}$ and G'_{V_2}. The subgraph G'_{V_2} is a complete graph with the set of vertices V_2 of the subgraph G_2. Since G_2 is a decomposable in the sense of Definition 1 graph with k cliques, according to the induction assumption, it can be obtained from G'_{V_2} by stepwise removing of edges such that each removed edge at the moment of its elimination belongs to exactly one clique. Hence G satisfies the condition in Definition 2. Consequently, regardless of the number of the cliques in a graph, if it is decomposable according to Definition 1 then it satisfies the condition in Definition 2.

Let now G be a decomposable according to Definition 2 graph. If G is a complete graph then it has only one clique and obviously the conditions of Definition 1 are satisfied. Let G be an incomplete graph which is obtained from a complete graph by stepwise removing of edges such that each removed edge at the moment of its elimination belongs to exactly one clique. Now we shall use Proposition 3. We begin with a complete graph, which is a decomposable graph in the sense of Definition 1, and each time we remove an edge which belongs to exactly one clique. Consequently we stay all the time in the set of graphs which are decomposable in the sense of Definition 1. □

P r o o f o f T h e o r e m 1. Using Lemma 1, Theorem 1 can be easily proved. Let the graph $G_{\_M}$ be decomposable. According to Definition 2, this graph can be obtained from a complete graph by stepwise removing of edges such that each removed edge at the moment of its elimination belongs to exactly one clique. The integration of the density function of $\mathbf{W}$ with respect to the variables from M can be done in the same sequence as the removing of the edges. Then the bounds of the integration for every element of M can be exactly obtained by directly applying Proposition 1 to the submatrix of $\mathbf{W}$ composed of the rows and columns with numbers from the set of vertices of the corresponding clique.

Assume now that the graph $G_{\_M}$ is non-decomposable. Consequently, for every sequence of removing of the excluded edges there always will be an edge which belongs to at least two cliques at the moment of its elimination. Each clique corresponds to a submatrix of the matrix $\mathbf{W}$. For every such submatrix Proposition 1 will give bounds a_i, b_i for the corresponding to this edge variable. The integration with respect to this variable have to be done in bounds from $\max(a_i)$ to $\min(b_i)$. $\square$

The proof of Lemma 1 shows a possible sequence for excluding of the chosen variables from the density of $\mathbf{W}$ such that the bounds of the integration with respect to every excluded variable can be exactly obtained by directly applying Proposition 1. Let the graph corresponding to the desired marginal density be decomposable with k cliques which are ordered as $C_1, \ldots, C_k$, according to Definition 1. Let us denote by V_i the set of vertices of C_i, $i = 1, \ldots, k$. We choose at first an arbitrary vertex v from the last vertex set V_k and exclude one after another all the variables with one of the indices v and the other index corresponding to a vertex outside the V_k. Then we repeat this procedure for another arbitrary chosen vertex from V_k and so on until we do this for all vertices in V_k. Next we repeat the same for the subgraph $C_1 \cup \ldots \cup C_{k-1}$ with respect to the last vertex set V_{k-1} in it and so on. Finally, we do this for the subgraph $C_1 \cup C_2$ with respect to its last vertex set V_2.

The quantity L in the marginal density given by Proposition 2 is more complicated when $\sigma^{q,r} \neq 0$. Using 8.445 in [2], L can be expressed in terms of infinite series as

$$L = \frac{\Gamma\left(\frac{n-p+1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n-p+2}{2}\right)} e^{-A\sigma^{q,r}} \left[1 + \sum_{k=1}^{\infty} \frac{(B\sigma^{q,r})^{2k}}{(\nu+1)\ldots(\nu+k)k!2^{2k}}\right],$$

where $\nu = \dfrac{n-p}{2}$. Hence, according to 9.14 1 in [2], L has a representation by the confluent hypergeometric limit functions ${}_0F_1(;\alpha;z)$ of the form

$$L = \frac{\Gamma\left(\frac{n-p+1}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n-p+2}{2}\right)} e^{-A\sigma^{q,r}} {}_0F_1\left(;\frac{n-p+2}{2};\left(\frac{B\sigma^{q,r}}{2}\right)^2\right).$$

When $\sigma^{q,r} = 0$, L takes the form (4).

Theorem 2. *Let* $\mathbf{W} = (W_{i,j})$ *has Wishart distribution* $\mathrm{W}_p(n, \Sigma)$ *and* q, r *be integers,* $1 \le q < r \le p$. *When* $\sigma^{q,r} = 0$, *the marginal density* $f_{\{W_{q,r}\}^c}(\mathrm{W}_0)$ *given by Proposition* 2, *can be written in the form*

$$f_{\{W_{q,r}\}^c}(\mathrm{W}_0) = \frac{f_{p-1,n,\Sigma[\{q\}^c]}(\mathrm{W}_0[\{q\}^c])\, f_{p-1,n,\Sigma[\{r\}^c]}(\mathrm{W}_0[\{r\}^c])}{f_{p-2,n,\Sigma[\{q,r\}^c]}(\mathrm{W}_0[\{q,r\}^c])}, \tag{7}$$

where $f_{p,n,\Sigma}(\mathrm{W})$ *is the Wishart density, given by* (2).

P r o o f. We shall prove the Theorem for $q = p-1$ and $r = p$ at first. From $\sigma^{p-1,p} = \sigma^{p,p-1} = 0$ it follows that $\det \Sigma[\{p-1\}^c, \{p\}^c] = \det \Sigma[\{p\}^c, \{p-1\}^c] = 0$. Using this fact and applying Proposition 4, we obtain the next equalities, which holds for $1 \le i, j < p-1$:

$$\det \Sigma \det \Sigma[\{i,p\}^c, \{j,p\}^c] = \det \Sigma[\{i\}^c, \{j\}^c] \det \Sigma[\{p\}^c] - \det \Sigma[\{i\}^c, \{p\}^c] \det \Sigma[\{p\}^c, \{j\}^c]; \tag{8}$$

$$\det \Sigma[\{p-1\}^c] \det \Sigma[\{i,p-1,p\}^c, \{j,p-1,p\}^c]$$

$$= \det \Sigma[\{i,p-1\}^c, \{j,p-1\}^c] \det \Sigma[\{p-1,p\}^c] - \det \Sigma[\{i,p-1\}^c, \{p-1,p\}^c] \det \Sigma[\{p-1,p\}^c, \{j,p-1\}^c]; \tag{9}$$

$$\det \Sigma \det \Sigma[\{i,p\}^c, \{p-1,p\}^c] = \det \Sigma[\{i\}^c, \{p-1\}^c] \det \Sigma[\{p\}^c]; \tag{10}$$

$$\det \Sigma \det \Sigma[\{p-1,p\}^c] = \det \Sigma[\{p-1\}^c] \det \Sigma[\{p\}^c]; \tag{11}$$

$$\det \Sigma \det \Sigma[\{i,p-1\}^c, \{p-1,p\}^c] = -\det \Sigma[\{i\}^c, \{p\}^c] \det \Sigma[\{p-1\}^c]. \tag{12}$$

We shall use the equalities (8), (9), (11), (12) to prove that

$$tr(\mathrm{W}_0\Sigma^{-1}) = tr\{\mathrm{W}_0[\{p-1\}^c](\Sigma[\{p-1\}^c])^{-1}\} + tr\{\mathrm{W}_0[\{p\}^c](\Sigma[\{p\}^c])^{-1}\} - tr\{\mathrm{W}_0[\{p-1,p\}^c](\Sigma[\{p-1,p\}^c])^{-1}\}. \tag{13}$$

Let us denote the elements of the matrices $(\Sigma[\{p-1\}^c])^{-1}$, $(\Sigma[\{p\}^c])^{-1}$ and $(\Sigma[\{p-1,p\}^c])^{-1}$ with $\sigma^{i,j}_{p-1}$, $\sigma^{i,j}_{p}$ and $\sigma^{i,j}_{p-1,p}$ respectively. Let $1 \le i, j < p-1$.

The coefficient of $w_{i,j}$ in the left hand side of (13) is $\sigma^{i,j}$, while in the right hand side it is $\sigma^{i,j}_{p-1} + \sigma^{i,j}_{p} - \sigma^{i,j}_{p-1,p}$. Consequently, we have to prove that

$$\frac{\det \Sigma[\{i\}^c, \{j\}^c]}{\det \Sigma} = \frac{\det \Sigma[\{i, p-1\}^c, \{j, p-1\}^c]}{\det \Sigma[\{p-1\}^c]} \tag{14}$$

$$+\frac{\det \Sigma[\{i, p\}^c, \{j, p\}^c]}{\det \Sigma[\{p\}^c]} - \frac{\det \Sigma[\{i, p-1, p\}^c, \{j, p-1, p\}^c]}{\det \Sigma[\{p-1, p\}^c]}.$$

Dividing (8) by $\det \Sigma \det \Sigma[\{p\}^c]$ we obtain that

$$\frac{\det \Sigma[\{i\}^c, \{j\}^c]}{\det \Sigma} - \frac{\det \Sigma[\{i, p\}^c, \{j, p\}^c]}{\det \Sigma[\{p\}^c]} = \frac{\det \Sigma[\{i\}^c, \{p\}^c] \det \Sigma[\{p\}^c, \{j\}^c]}{\det \Sigma \det \Sigma[\{p\}^c]}. \tag{15}$$

Dividing (9) by $\det \Sigma[\{p-1\}^c] \det \Sigma[\{p-1, p\}^c]$ we derive

$$\frac{\det \Sigma[\{i, p-1\}^c, \{j, p-1\}^c]}{\det \Sigma[\{p-1\}^c]} - \frac{\det \Sigma[\{i, p-1, p\}^c, \{j, p-1, p\}^c]}{\det \Sigma[\{p-1, p\}^c]} \tag{16}$$

$$= \frac{\det \Sigma[\{i, p-1\}^c, \{p-1, p\}^c] \det \Sigma[\{p-1, p\}^c, \{j, p-1\}^c]}{\det \Sigma[\{p-1\}^c] \det \Sigma[\{p-1, p\}^c]}.$$

From (15) and (16) we get that (14) is equivalent to the equality

$$\frac{\det \Sigma[\{i, \}^c, \{p\}^c] \det \Sigma[\{p\}^c, \{j\}^c]}{\det \Sigma \det \Sigma[\{p\}^c]}$$

$$= \frac{\det \Sigma[\{i, p-1\}^c, \{p-1, p\}^c] \det \Sigma[\{p-1, p\}^c, \{j, p-1\}^c]}{\det \Sigma[\{p-1\}^c] \det \Sigma[\{p-1, p\}^c]}$$

which can be verified using (11) and (12). By comparing the coefficients of $w_{i,p-1}$ in both sides of equality (13) we get the condition

$$\frac{\det \Sigma[\{i\}^c, \{p-1\}^c]}{\det \Sigma} = \frac{\det \Sigma[\{i, p\}^c, \{p-1, p\}^c]}{\det \Sigma[\{p\}^c]}$$

which follows directly from (10). Finally, comparing the coefficients of $w_{i,p}$ in both sides of (13) we get the relation

$$\frac{\det \Sigma[\{i\}^c, \{p\}^c]}{\det \Sigma} = \frac{-\det \Sigma[\{i, p-1\}^c, \{p-1, p\}^c]}{\det \Sigma[\{p-1\}^c]}$$

which follows from (12). Consequently, the representation (13) holds. The condition $1/\det\Sigma = (\det\Sigma[\{p-1,p\}^c])\,/\,(\det\Sigma[\{p-1\}^c]\det\Sigma[\{p\}^c])$ for the covariance matrix Σ can be derived from (11). Finally, we can easily check that $L/\Gamma_p(n/2) = \Gamma_{p-2}(n/2)/\left(\Gamma_{p-1}(n/2)\right)^2$, where L is given by (4). Consequently, the Theorem is true for $q = p-1$ and $r = p$.

Let now q, r be arbitrary integers, $1 \le q < r \le p$. Let in the matrix W_0 we place the q'th and r'th rows after the last row, and q'th and r'th columns after the last column. We shall obtain a matrix ${\mathrm{W}'}_0$ with $\det{\mathrm{W}'}_0 = \det\mathrm{W}_0$. Let us do the same with the matrix Σ and obtain analogously a matrix Σ' with $\det\Sigma' = \det\Sigma$. Since $tr(\mathrm{W}_0\Sigma^{-1}) = tr({\mathrm{W}'}_0\Sigma'^{-1})$, the density $f_{\{W_{q,r}\}^c}(\mathrm{W}_0)$ can be expresses in terms of ${\mathrm{W}'}_0$ and Σ'. The element $w_{q,r} = 0$ of W_0 lies on the $(p-1)$'th row and p'th column of the matrix ${\mathrm{W}'}_0$. The Theorem is true for $q = p-1$ and $r = p$, hence

$$f_{\{W_{q,r}\}^c}(\mathrm{W}_0) = \frac{f_{p-1,n,\Sigma'[\{p-1\}^c]}({\mathrm{W}'}_0[\{p-1\}^c])\; f_{p-1,n,\Sigma'[\{p\}^c]}({\mathrm{W}'}_0[\{p\}^c])}{f_{p-2,n,\Sigma'[\{p-1,p\}^c]}({\mathrm{W}'}_0[\{p-1,p\}^c])}$$

$$= \frac{f_{p-1,n,\Sigma[\{q\}^c]}(\mathrm{W}_0[\{q\}^c])\; f_{p-1,n,\Sigma[\{r\}^c]}(\mathrm{W}_0[\{r\}^c])}{f_{p-2,n,\Sigma[\{q,r\}^c]}(\mathrm{W}_0[\{q,r\}^c])}. \qquad \square$$

We shall denote the number of the elements of a set V by $|V|$.

Theorem 3. *Let $\mathbf{W} = (W_{i,j})$ has Wishart distribution $\mathrm{W}_p(n,\Sigma)$ and M be a subset of the set $\{W_{i,j}, 1 \le i < j \le p\}$ of non-diagonal elements of $\mathbf{W}$. Let $\sigma^{i,j} = 0$ for every element $W_{i,j}$ of M. Let the graph $G_{\_M}$ be decomposable with k cliques $C_1, \ldots, C_k$, ordered according to Definition 1 and V_i be the set of vertices of C_i, $i = 1, \ldots, k$. Then the joint density f_{M^c} of the elements of the set $M^c = \{W_{i,j}, 1 \le i < j \le p\}\backslash M$ can be written in the form*

$$f_{M^c}(\mathrm{W}_0) = \frac{\prod_{i=1}^{k} f_{|V_i|,n,\Sigma[\{V_i\}]}(\mathrm{W}_0[\{V_i\}])}{\prod_{i=2}^{k} f_{|U_i|,n,\Sigma[\{U_i\}]}(\mathrm{W}_0[\{U_i\}])},$$

where $\mathrm{W}_0 = (w_{i,j})$ is a $p \times p$ symmetric matrix, such that $w_{i,j} = 0$ for $W_{i,j} \in M$; $f_{p,n,\Sigma}(\mathrm{W})$ is the Wishart density function given by (2) and $U_i = (V_1 \cup \ldots \cup V_{i-1}) \cap V_i$, $i = 2, \ldots, k$.

P r o o f. The proof is by induction and follows the scheme of the proof of Lemma 1. □

REFERENCES

[1] T. W. ANDERSON. An Introduction to Multivariate Statistical Analysis. New York, John Wiley & Sons, 2nd ed., 2003.

[2] I. S. GRADSHTEYN, I. M. RYZHIK. Table of Integrals, Series, and Products. A. Jeffrey and D. Zwillinger, Eds. Elsevier, 7th ed., 2007.

[3] P. J. GREEN, A. THOMAS. Sampling decomposable graphs using a Markov chain on junction trees, *Biometrika*, 100, (2013), No 1, 91–110.

[4] S. L. LAURITZEN. Graphical Models. Oxford, Clarendon Press, 1996.

[5] R. J. MUIRHEAD. Aspects of Multivariate Statistical Theory. New York, John Wiley & Sons, 2nd ed., 2005.

[6] E. VELEVA. Some marginal densities of the Wishart distribution. *Math. and Education in Math.* **41** (2012), 273–278.

[7] E. VELEVA, T. IGNATOV. Distributions of joint sample correlation coefficients of independent normally distributed random variables. *Advanced Studies in Contemporary Mathematics* **12** (2006), No 2, 247–254.

[8] O.S. YAKOVLEV. A generalization of determinant identities (Russian). *Ukrain. Mat. Z.* **30** (1978), 850–854; English translation: Plenum Press, New York, 1979.

Evelina Veleva
Department of Applied Mathematics and Statistics
Ruse University "A. Kanchev"
8, Studentska Str.
7017 Ruse, Bulgaria
e-mail: `eveleva@abv.bg`