Доклади на Българската академия на науките
Comptes rendus de l'Académie bulgare des Sciences

Tome 59, No 4, 2006

MATHEMATIQUES
Théorie des probabilités

POSITIVE DEFINITE RANDOM MATRICES

E. Veleva

(*Submitted by Academician B. Boyanov on February 15, 2006*)

Abstract

The paper begins with necessary and sufficient conditions for positive definiteness of a matrix. These are used for obtaining an equivalent by distribution presentation of elements of a Wishart matrix by algebraic functions of independent random variables. Next section studies when sets of sample correlation coefficients of the multivariate normal (Gaussian) distribution are dependent, independent or conditionally independent. Some joint densities are obtained. The paper ends with an algorithm for generating uniformly distributed positive definite matrices with preliminary fixed diagonal elements.

Key words: positive definite matrix, Wishart distribution, multivariate normal (Gaussian) distribution, sample correlation coefficients, generating random matrices

2000 Mathematics Subject Classification: 62H10

1. Necessary and sufficient conditions for positive definiteness of a matrix.

Definition. Suppose $\mathbf{A}$ is a real symmetric matrix. If for any non-zero vector x, we have that

$$x'\mathbf{A}\,x > 0,$$

then $\mathbf{A}$ is a positive definite matrix (here and subsequently x' is the transpose of x).

Let $\mathbf{A} = \{a_{ij}\}_{i,j=1}^{n}$ be a real matrix, and let $N = \{1, 2, \ldots, n\}$. For any nonempty $\alpha, \beta \subset N$ let $\mathbf{A}(\alpha, \beta)$ be the submatrix of $\mathbf{A}$ obtained by deleting from $\mathbf{A}$ the rows indexed by α and the columns indexed by β. To simplify notation we use $\mathbf{A}(\alpha)$ instead of $\mathbf{A}(\alpha, \alpha)$. We will write the complement of α in N as α^c. For any integers i and j, $1 \le i \le j \le n$ we will denote by $\mathbf{A}(\alpha)_{ij}$ and $\mathbf{A}(\alpha)_{ij}^{0}$ the block matrices

$$\mathbf{A}(\alpha)_{ij} = \begin{pmatrix} \mathbf{A}(\alpha) & \mathbf{A}(\alpha, \{j\}^c) \\ \mathbf{A}'(\alpha, \{i\}^c) & a_{ij} \end{pmatrix}, \quad \mathbf{A}(\alpha)^0_{ij} = \begin{pmatrix} \mathbf{A}(\alpha) & \mathbf{A}(\alpha, \{j\}^c) \\ \mathbf{A}'(\alpha, \{i\}^c) & 0 \end{pmatrix}.$$

Theorem 1. Let $\mathbf{A}$ be a real symmetric matrix of size n, and let i,j be fixed integers such that $1 \leq i < j \leq n$. Matrix $\mathbf{A}$ is positive definite if and only if the matrices $\mathbf{A}(\{i\})$ and $\mathbf{A}(\{j\})$ are both positive definite and the element a_{ij} satisfies the inequalities

$$\frac{-\left|\mathbf{A}(\{i,j\})^0_{ij}\right| - \sqrt{|\mathbf{A}(\{i\})|\ |\mathbf{A}(\{j\})|}}{|\mathbf{A}(\{i,j\})|} < a_{ij} < \frac{-\left|\mathbf{A}(\{i,j\})^0_{ij}\right| + \sqrt{|\mathbf{A}(\{i\})|\ |\mathbf{A}(\{j\})|}}{|\mathbf{A}(\{i,j\})|},$$

where $|\,.\,|$ denotes determinant of a matrix.

Theorem 2. A real symmetric matrix $\mathbf{A} = \{a_{ij}\}^n_{i,j=1}$ is positive definite if and only if its elements satisfy the following conditions:

$$a_{ii} > 0, \quad i = 1, \ldots, n;$$

$$-\sqrt{a_{11}\, a_{ii}} < a_{1i} < \sqrt{a_{11}\, a_{ii}}, \quad i = 2, \ldots, n;$$

$$\frac{-\left|\mathbf{A}(\{i,\ldots,n\})^0_{ij}\right| - \sqrt{|\mathbf{A}(\{i,\ldots,n\})_{ii}|\,|\mathbf{A}(\{i,\ldots,n\})_{jj}|}}{|\mathbf{A}(\{i,\ldots,n\})|} < a_{ij}$$

$$< \frac{-\left|\mathbf{A}(\{i,\ldots,n\})^0_{ij}\right| + \sqrt{|\mathbf{A}(\{i,\ldots,n\})_{ii}|\,|\mathbf{A}(\{i,\ldots,n\})_{jj}|}}{|\mathbf{A}(\{i,\ldots,n\})|},$$

$$i = 2, \ldots, n-1, \quad j = i+1, \ldots, n.$$

The proofs of Theorems 1 and 2 will appear in [1].

2. Wishart random matrices. Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be independent, $\mathbf{X}_i \sim N_p(\mathbf{0}, \mathbf{\Sigma})$, with $n \geq p$, i.e., $\mathbf{X}_i$ has a p-dimensional multivariate normal distribution with mean vector a zero vector $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma}$, $\mathbf{\Sigma} > 0$ and let $\mathbf{W} = \sum \mathbf{X}_i \mathbf{X}'_i$. We say that $\mathbf{W}$ has a central Wishart distribution with n degrees of freedom and covariance matrix $\mathbf{\Sigma}$ and write $\mathbf{W} \sim W_p(n, \mathbf{\Sigma})$. It is easy to see that $\mathbf{W}$ is a $p \times p$ matrix and that $\mathbf{W} > 0$. From this definition it is apparent that

$$E\,\mathbf{W} = n\,\mathbf{\Sigma},$$

$$\mathbf{AWA}' \sim W_q(n, \mathbf{A\Sigma A}'), \tag{1}$$

where $\mathbf{A}$ is $q \times p$ of rank q.

It is known that for an arbitrary positive definite $p \times p$ matrix $\mathbf{\Sigma}$ there exists a $p \times p$ matrix $\mathbf{B}$ such that $\mathbf{\Sigma} = \mathbf{BB}'$. Consequently, from (1) it follows that if $\mathbf{W} \sim W_p(n, \mathbf{I})$, where $\mathbf{I}$ is the identity matrix of size p, then $\mathbf{BWB}' \sim W_p(n, \mathbf{\Sigma})$. Therefore, from now on we consider the case $\mathbf{\Sigma} = \mathbf{I}$.

IGNATOV and NIKOL'OVA in [2] consider the distribution $\Psi(n,p)$ with joint density

$$C \mid \mathbf{D} \mid^{(n-p-1)/2} I_E, \tag{2}$$

where

- C is a constant;
- $\mathbf{D}$ is the matrix

$$\mathbf{D} = \begin{pmatrix} 1 & x_{12} & \dots & x_{1p} \\ x_{12} & 1 & \dots & x_{2p} \\ \cdot & \cdot & \cdot & \cdot \\ x_{1p} & x_{2p} & \dots & 1 \end{pmatrix};$$

- I_E is the indicator of the set E containing all points $(x_{ij},\ 1 \le i < j \le p)$ in the real space $R_{p(p-1)/2}$, such that the matrix $\mathbf{D}$ is positive definite.

They have proved the following Proposition:

Proposition. Let $\tau_1, \tau_2, \dots, \tau_p$ be independent and identically χ^2-distributed random variables with n degrees of freedom and let random variables $\nu_{ij},\ 1 \le i < j \le p$ have distribution $\Psi(n,p)$. Let us assume that the set $\{\tau_1, \tau_2, \dots, \tau_p\}$ is independent of the set $\{\nu_{ij},\ 1 \le i < j \le p\}$. Then the matrix below $\boldsymbol{V}$ has the Wishart distribution $W_p(n, \mathbf{I})$

$$\boldsymbol{V} = \begin{pmatrix} \sqrt{\tau_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sqrt{\tau_p} \end{pmatrix} \begin{pmatrix} 1 & \nu_{12} & \dots & \nu_{1p} \\ \nu_{12} & 1 & \dots & \nu_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \nu_{1p} & \nu_{2p} & \dots & 1 \end{pmatrix} \begin{pmatrix} \sqrt{\tau_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sqrt{\tau_p} \end{pmatrix}.$$

Using the same variable transformation as in the proof of this Proposition, it can be shown that the opposite statement is also true. So, the density (2) gives actually the joint density of the entries of the sample correlation matrix in the case $\mathbf{\Sigma} = \mathbf{I}$.

The next Theorem gives an equivalent by distribution presentation of the random variables $\nu_{ij},\ 1 \le i < j \le p$, having distribution $\Psi(n,p)$ by functions of independent random variables. This can be used for generating $\Psi(n,p)$ distributed random matrices and consequently Wishart ones (see [3]).

Theorem 3. Let the random variables $\eta_{ij},\ 1 \le i < j \le p$ be mutually independent. Suppose that $\eta_{i\,i+1}, \dots, \eta_{ip}$ be identically distributed $\Psi(n-i+1, 2)$ for $i = 1, \dots, p-1$. The random variables $\nu_{ij},\ 1 \le i < j \le p$, defined by

$$\nu_{1i} = \eta_{1i}, \quad i = 2, \dots, p,$$

$$\nu_{2i} = \eta_{12}\eta_{1i} + \eta_{2i}\sqrt{(1-\eta_{12}^2)(1-\eta_{1i}^2)}, \quad i = 3, \ldots, p,$$

$$\nu_{ij} = \eta_{ij} \prod_{q=1}^{i-1} \sqrt{(1-\eta_{qi}^2)(1-\eta_{qj}^2)} + \sum_{t=1}^{i-1} \left(\eta_{ti}\, \eta_{tj} \prod_{q=1}^{t-1} \sqrt{(1-\eta_{qi}^2)(1-\eta_{qj}^2)} \right), \quad 3 \le i < j \le p$$

have distribution $\Psi(n,p)$.

To generate random $\Psi(n,p)$ matrices using Theorem 3, we have to know how to generate random variables with distribution $\Psi(k,2)$ for $k \ge 2$. For $p = 2$ the density (2) of the distribution $\Psi(n,p)$ has the form

$$C\,(1-y^2)^{\frac{n-3}{2}}, \quad y \in (-1,\, 1)$$

and it is known as the Pearson probability distribution of second type. Random variables with this distribution can be easily generated (see [4], p. 481) using the quotient of the difference and the sum of two gamma-distributed random variables.

3. Marginal densities of the distribution $\Psi(n,p)$. In recent years the Wishart distribution and distributions derived from the Wishart one have received a lot of attention because of their use in graphical Gaussian models. The essence of graphical models in multivariate analysis is to identify independences and conditional independences between various groups of variables.

Empirical correlation matrices are of great importance for risk management and asset allocation. The study of correlation matrices is one of the cornerstone of Markowitz's theory of optimal portfolios.

Using Theorems 1 and 2 it is possible to get some marginal densities of the distribution $\Psi(n,p)$ and to investigate when sets of sample correlation coefficients of the multivariate normal (Gaussian) distribution are dependent, independent or conditionally independent.

Let ν_{ij}, $1 \le i < j \le p$ be random variables with distribution $\Psi(n,p)$. Denote by V the set of random variables $V = \{\nu_{ij},\ 1 \le i < j \le p\}$. In [5] we have proved the following Theorem:

Theorem 4. Every p random variables from the set V are dependent.

Let us introduce the notation f_S for the joint density of the random variables from a set S; if $S = \Phi$, i.e. the set S is empty we define $f_S = 1$.

Theorem 5. Let r and q be arbitrary integers, such that $2 \le r \le p-1$ and $1 \le q \le p-2$. Let us denote by S_1 and S_2 the sets $S_1 = \{\nu_{ij} \mid 1 \le i < j \le r\}$ and $S_2 = \{\nu_{ij} \mid q+1 \le i < j \le p\}$. Then

$$f_{S_1 \cup S_2} = \frac{f_{S_1}\, f_{S_2}}{f_{S_1 \cap S_2}}. \tag{3}$$

The proof will appear in [6].

We denote by $f_{S_1 \cup S_2 / S_1 \cap S_2}$, $f_{S_1/S_1 \cap S_2}$ and $f_{S_2/S_1 \cap S_2}$ the densities of $S_1 \cup S_2$, S_1 and S_2, conditioned on the random variables from the set $S_1 \cap S_2$. If we divide the two sides of equality (3) by $f_{S_1 \cap S_2}$ we get the relation

$$f_{S_1 \cup S_2 / S_1 \cap S_2} = f_{S_1/S_1 \cap S_2}\, f_{S_2/S_1 \cap S_2}. \tag{4}$$

According to the definition of conditional independence given in [7], equality (4) shows that the random variables from the sets S_1 and S_2 are independent conditionally on the random variables from the set $S_1 \cap S_2$.

Theorem 6. Let $q_1, \ldots, q_k, r_1, \ldots r_k$ be integers such that $1 = q_1 \leq q_2 \leq \ldots \leq q_k \leq p$ and $1 \leq r_1 \leq r_2 \leq \ldots \leq r_k = p$. Let us denote by S_l the set $S_l = \{v_{ij} \mid q_l \leq i < j \leq r_l\}$, $l = 1, \ldots k$. Then

$$f_{S_1 \cup \ldots \cup S_k} = \frac{f_{S_1} f_{S_2} \cdots f_{S_k}}{f_{S_1 \cap S_2} f_{S_2 \cap S_3} \cdots f_{S_{k-1} \cap S_k}}.$$

Let $S = \{\nu_{i_s j_s}, s = 1, \ldots, k\}$ be an arbitrary subset of the set of random variables V. To this subset S we can attach a graph $G(S)$ with nodes $\{1, 2, \ldots, p\}$ and k undirected edges $\{\ \{i_s, j_s\},\ s = 1, \ldots, k\}$.

Theorem 7. Let S_1 and S_2 be two subsets of the set V and G_1 and G_2 be their corresponding graphs. Let us denote by K_1 the set of the numbers of the nodes which are vertex of an edge from G_1. Analogically, for graph G_2 let us denote the corresponding set by K_2 and let $K = K_1 \cap K_2$. If the set K contains at most one element then

$$f_{S_1 \cup S_2} = f_{S_1} f_{S_2},$$

i.e. the set S_1 and S_2 are independent. If the set K contains at least two elements and for every two elements r and q from K, $r < q$ the random variable ν_{rq} belongs simultaneously to S_1 and S_2 then

$$f_{S_1 \cup S_2} = \frac{f_{S_1} f_{S_2}}{f_{S_1 \cap S_2}},$$

i.e. the sets S_1 and S_2 are independent conditioned on the set $S_1 \cap S_2$.

Theorem 8. Let S be a subset of the set V and G be the corresponding graph. Let us denote by S_+ the subset of S containing all variables which edges in the graph G are part of a closed path (circuit). If S_- is the set $S_- = S \backslash S_+$ then the sets S_+ and S_- are independent.

The proofs of Theorems 6 – 8 will appear in [8].

In [9] we have proved the following two Theorems:

Theorem 9. Let r and q be arbitrary integers such that $1 \leq r < q \leq p$. The random variables from the set $S = V \backslash \{\nu_{rq}\}$ have joint density function of the form

$$C\ \frac{\Gamma\left(\frac{(n-p+1)}{2}\right)\ \Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{n-p+2}{2}\right)}\ \frac{\left(\ |\mathbf{D}_{r;r}|\ |\mathbf{D}_{q;q}|\right)^{\frac{(n-p)}{2}}}{|\mathbf{D}_{r,q;r,q}|^{\frac{(n-p+1)}{2}}}\ I_E,$$

where

- I_E is the indicator of the set E containing all points in $R_{p(p-1)/2-1}$ such that the matrices $\mathbf{D}_{r;r}$ and $\mathbf{D}_{q;q}$ are both positive definite;
- $\Gamma(.)$ is the well-known Gamma function.

Theorem 10. Let r be an arbitrary integer such that $1 \le r \le p$. The random variables from the set $S = \{\nu_{ij} \,|1 \le i < j \le p,\ i \ne r,\ j \ne r\}$ have joint density function of the form

$$f_S = \frac{\left[\Gamma\left(\frac{n}{2}\right)\right]^{p-2}}{\Gamma\left(\frac{n-1}{2}\right)\Gamma\left(\frac{n-2}{2}\right)\ldots\Gamma\left(\frac{n-p+2}{2}\right)\left[\Gamma\left(\frac{1}{2}\right)\right]^{\frac{(p-1)(p-2)}{2}}}\,|\mathbf{D}_{r;r}|^{\frac{n-p}{2}}$$

for all points in $R_{(p-1)(p-2)/2}$ for which the matrix $\mathbf{D}_{r;r}$ is positive definite.

4. Uniformly distributed positive definite matrices. The positive definiteness of a matrix is a necessary condition for applying many numerical algorithms. The diagonal elements of the matrix have often specified significance. The correctness of such numerical algorithm can be proven if we are able to choose a positive definite matrix at random with uniform distribution. The space of all positive definite matrices is, however, a cone and consequently a uniform distribution cannot be defined over the whole cone because it has infinite volume. This enforces to introduce additional restrictions on the matrices, which together with the positive definiteness reduce our choice within a set with finite volume. An algorithm for generating uniformly distributed positive definite matrices with fixed diagonal elements is given below. If the user does not want to fix concrete diagonal elements of the matrix in advance, he has though to assign bounds for each diagonal element. For instance, he can choose all diagonal elements to be in the interval (0,100). Then a uniform random number in the chosen bounds has to be generated for every diagonal element. The algorithm described below allows the user a random choice among all positive definite matrices with concrete diagonal elements that are either fixed in advance or randomly generated within the chosen bounds.

From now on we assume that $a_{11}, \ldots, a_{pp}$ are the diagonal elements of the matrix, chosen in accordance with users' preferences. They have to be positive so that to exist at least one positive definite matrix with such diagonal elements (see Theorem 1).

Theorem 11. Let the random variables $\eta_{ij},\ 1 \le i < j \le p$ be mutually independent. Suppose that $\eta_{i\,i+1}, \ldots, \eta_{ip}$ be identically distributed $\Psi(p-i+2, 2)$ for $i = 1, \ldots, p-1$. Consider the random variables $\nu_{ij},\ 1 \le i < j \le p$, defined by

$$\nu_{1i} = \eta_{1i}\sqrt{a_{11}a_{ii}}, \quad i = 2, \ldots, p,$$

$$\nu_{ij} = \sqrt{a_{ii}a_{jj}}\left[\sum_{r=1}^{i-1}\left(\eta_{ri}\eta_{rj}\prod_{q=1}^{r-1}\sqrt{(1-\eta_{qi}^2)(1-\eta_{qj}^2)}\right) + \eta_{ij}\prod_{q=1}^{i-1}\sqrt{(1-\eta_{qi}^2)(1-\eta_{qj}^2)}\right]$$

$$i = 2, \ldots, p-1, \quad j = i+1, \ldots, p.$$

The joint density function of the random variables $\nu_{ij},\ 1 \le i < j \le p$ is of the form

$$f_{\nu_{ij}, 1 \le i < j \le p}(x_{ij}, 1 \le i < j \le p) = CI_E,$$

where

• C is the constant

$$C = \frac{\left[\Gamma\left(\frac{p+1}{2}\right)\right]^{p-1}}{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{p-1}{2}\right)\ldots\Gamma\left(\frac{3}{2}\right)\left[\Gamma\left(\frac{1}{2}\right)\right]^{\frac{p(p-1)}{2}}(a_{11}a_{22}\ldots a_{pp})^{\frac{p-1}{2}}};$$

• I_E is the indicator of the set E consisting of all points $(x_{ij}, 1 \leq i < j \leq p)$ in $R_{p(p-1)/2}$ for which the matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & x_{12} & \cdots & x_{1p} \\ x_{12} & a_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{1p} & x_{2p} & \cdots & a_{pp} \end{pmatrix} \tag{5}$$

is positive definite.

Theorem 11 gives the following algorithm for generating uniformly distributed positive definite matrices:

1) Generate $p(p-1)/2$ random numbers y_{ij}, $1 \leq i < j \leq p$ so that y_{ij} comes from the distribution $\Psi(p-i+2, 2)$.

2) In order to reduce calculations, compute the auxiliary quantities z_{ij}, $1 \leq i \leq j \leq p$ such that

$$z_{ij} = y_{ij}\sqrt{(1-y_{1j}^2)\ldots(1-y_{i-1j}^2)} \text{ with } y_{ii} = 1.$$

3) Calculate the desired matrix (5) with

$$x_{ij} = \sqrt{a_{ii}a_{jj}}(z_{1i}z_{1j} + z_{2i}z_{2j} + \cdots + z_{ii}z_{ij}).$$

The obtained formulas show that it is possible to create a programme which generates without dialogue with the user a matrix $\mathbf{U}$ with units on the main diagonal. This is done by steps 1)–3) substituting $a_{11} = \cdots = a_{pp} = 1$. Then the user according to his preferences forms a diagonal matrix $\mathbf{D} = \operatorname{diag}(\sqrt{a_{11}}, \ldots, \sqrt{a_{pp}})$ and the desired matrix $\mathbf{A}$ is $\mathbf{A} = \mathbf{DUD}$.

The details will appear in [10].

REFERENCES

[1] VELEVA E. Linear Algebra and Appl. (submitted). [2] IGNATOV T. G., A. D. NIKOLOVA. Annuaire Univ. Sofia, Fac. of Econ. & Buss. Admin., **3**, 2004, 79–94. [3] VELEVA E. Ibid., **6**, 2006 (to appear). [4] DEVROYE L. Non-Uniform Random Variate Generation, New York, Springer-Verlag, 1986, 843. [5] VELEVA E., T. IGNATOV. Advanced Studies in Contemporary Mathematics, **12**, 2006, No 2 (to appear). [6] VELEVA E. In: Proceedings of 35th Spring Conference of the Union of Bulgarian Mathematicians, 5–8 April, 2006, Borovec, 315–320. [7] DAWID A. P., J. ROY. Stat. Soc., Ser. B, **41**, 1979, No 1, 1–31. [8] VELEVA E. Joint densities of correlation coefficients for samples from multivariate standard normal

distribution (in preparation). [9] Id. Annuaire Univ. Sofia, Fac. of Econ. & Buss. Admin., **5**, 2005 (to appear). [10] Id. Proceedings of 35th Spring Conference of the Union of Bulgarian Mathematicians, 5 – 8 April, 2006, Borovec, 321–326.

Department of Numerical Methods and Statistics
Rousse University
8, Studentska Str.
7017 Rousse, Bulgaria
e-mail: `eveleva@abv.bg`