V. S. Mukha, N. F. Kako

Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus

INTEGRALS AND INTEGRAL TRANSFORMATIONS RELATED TO THE VECTOR GAUSSIAN DISTRIBUTION

Abstract. This paper is dedicated to the integrals and integral transformations related to the probability density function of the vector Gaussian distribution and arising in probability applications. Herein, we present three integrals that permit to calculate the moments of the multivariate Gaussian distribution. Moreover, the total probability formula and Bayes formula for the vector Gaussian distribution are given. The obtained results are proven. The deduction of the integrals is performed on the basis of the Gauss elimination method. The total probability formula and Bayes formula are obtained on the basis of the proven integrals. These integrals and integral transformations could be used, for example, in the statistical decision theory, particularly, in the dual control theory, and as table integrals in various areas of research. On the basis of the obtained results, Bayesian estimations of the coefficients of the multiple regression function are calculated.

Keywords: vector Gaussian distribution, multidimensional integrals, total probability formula, Bayes formula, multiple regression function, Bayesian estimations

В. С. Муха, Н. Ф. Како

Белорусский государственный университет информатики и радиоэлектроники, Минск, Беларусь

ИНТЕГРАЛЫ И ИНТЕГРАЛЬНЫЕ ПРЕОБРАЗОВАНИЯ, СВЯЗАННЫЕ С ВЕКТОРНЫМ ГАУССОВСКИМ РАСПРЕДЕЛЕНИЕМ

Аннотация. Рассматриваются интегралы и интегральные преобразования, относящиеся к функции плотности вероятности векторного гауссовского распределения и возникающие в вероятностных приложениях. Представлены три интеграла, позволяющие рассчитывать моменты векторного гауссовского распределения, а также формулы полной вероятности и Байеса. Приводятся доказательства полученных результатов. Вывод интегралов выполнен на основе метода исключения Гаусса. Формулы полной вероятности и Байеса получены на основе доказанных интегралов. Представленные интегралы и интегральные преобразования могут быть использованы в различных вероятностных приложениях, например в теории статистических решений, в частности, в теории дуального управления, а также как табличные интегралы в различных областях исследований. На основе полученных результатов рассчитаны байесовские оценки коэффициентов множественной функции регрессии.

Ключевые слова: векторное гауссовское распределение, многомерные интегралы, формула полной вероятности, формула Байеса, множественная функция регрессии, байесовские оценки

Introduction. Integrals and integral transformations related to probability distributions have many applications, one of them being the statistical decision theory. The latter attracts much attention due to its capacity to formulate problems in a strict mathematical form. One of the technical problems solved by the statistical decision theory is the problem of dual control [1], requiring the calculation of integrals related to multivariate probability distributions. In this paper, we present three integrals connected with the vector Gaussian distribution and the total probability formula and Bayes formula for the vector Gaussian distribution.

1. Integrals connected with the vector Gaussian distribution. A random vector with k components $\Xi^T = (\Xi_1, \Xi_2, ..., \Xi_k)$ is distributed according to the normal or Gaussian law if its probability density has the form

$$f(\xi) = \frac{1}{\sqrt{(2\pi)^k |d_\Xi|}} \exp\left(-\frac{1}{2}(\xi - \nu_\Xi)^T d_\Xi^{-1}(\xi - \nu_\Xi)\right), \quad \xi \in E^k, \tag{1}$$

where $\xi^T = (\xi_1, \xi_2, ..., \xi_k)$ is the vector-row of the arguments of the probability density $f(\xi)$, $\nu_\Xi^T = (\nu_{\Xi,1}, \nu_{\Xi,2}, ..., \nu_{\Xi,k})$ is the vector-row of the parameters of the probability density $f(\xi)$, $d_\Xi = (d_{\Xi,i,j})$, $i, j = \overline{1,k}$, is the symmetric positive-definite matrix of the parameters of the probability density $f(\xi)$, d_Ξ^{-1} is the matrix inverse to the matrix d_Ξ, $|d_\Xi|$ is the determinant of the matrix d_Ξ, E^k is the k-dimensional Euclidean space, and the T symbol stands for *transpose*. The parameters ν_Ξ and d_Ξ of distribution (1) are mathematical expectation and dispersion (variance-covariance) matrices of the random vector Ξ, respectively [2].

The following equalities for function (1) are true:

$$\int_{E^k} \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right)d\xi = \sqrt{(2\pi)^k |A^{-1}|}\exp\left(\frac{1}{2}B^T A^{-1}B\right), \tag{2}$$

$$\int_{E^k} C^T\xi \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right)d\xi = \sqrt{(2\pi)^k |A^{-1}|}\exp\left(\frac{1}{2}B^T A^{-1}B\right)C^T A^{-1}B, \tag{3}$$

$$\int_{E^k} \xi^T U\xi \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right)d\xi = \sqrt{(2\pi)^k |A^{-1}|}\left(\mathrm{Tr}(A^{-1}U) + B^T A^{-1}UA^{-1}B\right)\exp\left(\frac{1}{2}B^T A^{-1}B\right), \tag{4}$$

where $A = (a_{i,j})$, $i, j = \overline{1,k}$, is the symmetric positive-definite matrix, $\xi^T = (\xi_1, \xi_2, ..., \xi_k)$, $B^T = (b_1, b_2, ..., b_k)$, $C^T = (c_1, c_2, ..., c_k)$ are the vector-rows, A^{-1} is the matrix inverse to the matrix A, $|A^{-1}|$ is the determinant of the matrix A^{-1}, $U = (u_{i,j})$, $i, j = \overline{1,k}$, is the symmetric positive-definite matrix, and $\mathrm{Tr}(A^{-1}U)$ is the trace of the matrix $A^{-1}U$.

Integrals (2)–(4) were received in [3], but proof was given only for equality (2). Now we give more detailed proof of equality (2) and prove equalities (3) and (4).

P r o o f   o f   e q u a l i t y   (2). We will use the Gauss elimination method [4], which permits to bring the matrix A to the diagonal form and reduce the calculation of the multiple integral to the calculation of the repeated integral. The conditions for using the Gauss elimination method in this case are held because the matrix A is positive-definite.

We denote the integrand function $F(\xi)$ in (2):

$$F(\xi) = \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right). \tag{5}$$

By applying the Gauss elimination method [4] to the matrix $A = (a_{i,j})$ in (5), we receive the upper triangular matrix

$$G = \begin{pmatrix} a_{1,1}^{(0)} & a_{1,2}^{(0)} & \cdots & a_{1,k}^{(0)} \\ 0 & a_{2,2}^{(1)} & \cdots & a_{2,k}^{(1)} \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & a_{k,k}^{(k-1)} \end{pmatrix} = \left(a_{i,j}^{(i-1)}\right), \quad i, j = \overline{1,k}.$$

The determinant of the upper triangular matrix G is equal to the product of its diagonal elements and also to the determinant of the matrix A [4]:

$$|G| = a_{1,1}^{(0)} a_{2,2}^{(1)} \cdots a_{k,k}^{(k-1)} = |A| = \prod_{i=1}^{k} a_{i,i}^{(i-1)}.$$

The mine minors of the matrices A and G are equal, $A_m = G_m$, $m = 1,2,...,k$ [4], where $A_m = \begin{pmatrix} 1,2,...,m \\ 1,2,...,m \end{pmatrix}$ is the main minor of order m of the matrix A and $G_m = \begin{pmatrix} 1,2,...,m \\ 1,2,...,m \end{pmatrix}$ is the main minor of order m of the matrix G. Moreover [4],

$$a_{1,1}^{(0)} = A_1, \quad a_{2,2}^{(1)} = \frac{A_2}{A_1}, \quad a_{3,3}^{(2)} = \frac{A_3}{A_2},..., \quad a_{k,k}^{(k-1)} = \frac{A_k}{A_{k-1}}. \tag{6}$$

As shown in [4], the matrix A can be represented in the form

$$A = G_l^T \widehat{D}_1 G, \tag{7}$$

where G and G_l are matrices received by the Gauss elimination method from the matrices A and A^T, respectively, and $\widehat{D}_1$ is the diagonal matrix represented in the form

$$\widehat{D}_1 = \text{diag} \left\{ \frac{1}{A_1}, \frac{A_1}{A_2}, \frac{A_2}{A_3},..., \frac{A_{k-1}}{A_k} \right\}.$$

Since the matrix A is symmetrical, i. e. $A^T = A$, then $G_l = G$, and equation (7) can be written in the form:

$$A = G^T \widehat{D}_1 G.$$

Having expression (6), we get that the matrix A can be represented in the following form:

$$A = G^T \widehat{D} G, \tag{8}$$

where $\widehat{D}$ is a diagonal matrix

$$\widehat{D} = \text{diag} \left\{ \left(a_{1,1}^{(0)} \right)^{-1}, \left(a_{2,2}^{(1)} \right)^{-1},..., \left(a_{k,k}^{(k-1)} \right)^{-1} \right\}. \tag{9}$$

From (8) it follows that

$$\widehat{D}^{-1} = G A^{-1} G^T. \tag{10}$$

We denote as $\widehat{d}_{i,i} = (a_{i,i}^{(i-1)})^{-1}$, $i = \overline{1,k}$, the diagonal elements of the diagonal matrix $\widehat{D}$ (9). Meanwhile,

$$|\widehat{D}^{-1}| = |G| = |A| = \prod_{i=1}^{k} a_{i,i}^{(i-1)}. \tag{11}$$

Let us introduce in (5) the linear replacement of the variables

$$\xi = G^{-1} z, \tag{12}$$

with which the integrand function (5) is converted to the following function of the z argument:

$$F(z) = \exp\left(-\frac{1}{2}z^T P z + Dz\right),$$

where

$$P = (G^{-1})^T A G^{-1},$$

$$D = B^T G^{-1} = (d_i), \quad i = \overline{1,k}. \tag{13}$$

Since, provided (8),

$$P = (G^{-1})^T A G^{-1} = (G^{-1})^T G^T \hat{D} G G^{-1} = \hat{D},$$

we have the following function of the z argument:

$$F(z) = \exp\left(-\frac{1}{2}z^T \hat{D} z + Dz\right). \tag{14}$$

The following equality is obtained after replacing the variables in integral [5]:

$$\int\limits_{E^k} F(\xi)d\xi = \int\limits_{E^k} F(z)\,|\,J\,|\,dz, \tag{15}$$

where $|\,J\,|$ is the absolute value of the Jacobian of transformation (12):

$$|\,J\,| = |\,G^{-1}\,| = |\,\hat{D}\,| = \prod_{i=1}^{k}(a_{i,i}^{(i-1)})^{-1}.$$

Let us rewrite the function $F(z)$ in (14) as a function of the elements of the matrices $\hat{D}$ and D, taking into account the notations above:

$$F(z) = \exp\left(-\frac{1}{2}z^T \hat{D} z + Dz\right) = \exp\left(-\frac{1}{2}\sum_{i=1}^{k}(a_{i,i}^{(i-1)})^{-1}z_i^2 + \sum_{i=1}^{k}d_i z_i\right) = \prod_{i=1}^{k}\exp\left(-\frac{1}{2}(a_{i,i}^{(i-1)})^{-1}z_i^2 + d_i z_i\right). \tag{16}$$

Substituting (16) into (15) we obtain the following equality:

$$\int\limits_{E^k} F(\xi)d\xi = |\,J\,|\prod_{i=1}^{k}\int\limits_{-\infty}^{\infty}\exp\left(-\frac{1}{2}(a_{i,i}^{(i-1)})^{-1}z_i^2 + d_i z_i\right)dz_i. \tag{17}$$

The integral in the right part of expression (17) is a table integral [6]:

$$\int\limits_{-\infty}^{\infty}\exp\left(-\frac{1}{2}\alpha x^2 + \beta x\right)dx = \sqrt{\frac{2\pi}{\alpha}}\exp\left(\frac{\beta^2}{2\alpha}\right). \tag{18}$$

Taking into account integral (18) we receive instead of (17):

$$\int\limits_{E^k} F(\xi)d\xi = |\,J\,|\prod_{i=1}^{k}\sqrt{2\pi a_{i,i}^{(i-1)}}\exp\left(-\frac{1}{2}\sum_{i=1}^{k}a_{i,i}^{(i-1)}d_i^2\right),$$

or in the matrix form

$$\int_{E^k} F(\xi)d\xi = \sqrt{(2\pi)^k |\widehat{D}|}\exp\left(-\frac{1}{2}D\widehat{D}^{-1}D^T\right). \tag{19}$$

Let us turn back from the matrices D and $\widehat{D}$ to the matrices A and B. Since $|\widehat{D}|=|A^{-1}|$ (formula (11)), $\widehat{D}^{-1} = GA^{-1}G^T$ (formula (10)), and $D = B^T G^{-1}$ (formula (13)), then

$$D\widehat{D}^{-1}D^T = B^T G^{-1}GA^{-1}G^T(G^{-1})^T B = B^T A^{-1}B,$$

and instead of (19) we receive (2). Equality (2) is proven.

P r o o f o f e q u a l i t y (3). We calculate the integral

$$I_1 = \frac{1}{\sqrt{(2\pi)^k |A^{-1}|}}\int_{E^k} C^T\xi\exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right)d\xi. \tag{20}$$

Let us complete the square of the expression $-\xi^T A\xi/2 + B^T\xi$ in integral (20). We receive the expression

$$-\frac{1}{2}\xi^T A\xi + B^T\xi = -\frac{1}{2}(\xi - A^{-1}B)^T A(\xi - A^{-1}B) + \frac{1}{2}B^T A^{-1}B, \tag{21}$$

and instead of integral (20) we receive the following integral:

$$I_1 = \frac{1}{\sqrt{(2\pi)^k |A^{-1}|}}\int_{E^k} C^T\xi\exp\left(-\frac{1}{2}(\xi - A^{-1}B)^T A(\xi - A^{-1}B)\right)\exp\left(\frac{1}{2}B^T A^{-1}B\right)d\xi. \tag{22}$$

Since the function

$$f(\xi) = \frac{1}{\sqrt{(2\pi)^k |A^{-1}|}}\exp\left(-\frac{1}{2}(\xi - A^{-1}B)^T A(\xi - A^{-1}B)\right) \tag{23}$$

in (22) is the probability density of the vector Gaussian distribution with the mean value $A^{-1}B$ and the dispersion matrix A^{-1}, then:

$$I_1 = \frac{1}{\sqrt{(2\pi)^k |A^{-1}|}}\int_{E^k} C^T\xi\exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right)d\xi = \exp\left(\frac{1}{2}B^T A^{-1}B\right)C^T A^{-1}B.$$

As a result, we have equality (3).

P r o o f o f e q u a l i t y (4). We will perform it in a way similar to the proof of equality (3). We calculate the integral

$$I_2 = \frac{1}{\sqrt{(2\pi)^k |A^{-1}|}}\int_{E^k} \xi^T U\xi\exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi\right)d\xi. \tag{24}$$

Completing the square of the expression $-\xi^T A\xi/2 + B^T\xi$ in integral (24) gives expression (21). Thus, instead of integral (21) we have:

$$I_2 = \frac{1}{\sqrt{(2\pi)^k |A^{-1}|}}\int_{E^k} \xi^T U\xi\exp\left(-\frac{1}{2}(\xi - A^{-1}B)^T A(\xi - A^{-1}B)\right)\exp\left(\frac{1}{2}B^T A^{-1}B\right)d\xi. \tag{25}$$

Since the function in (25) of form (23) is the probability density of the vector Gaussian distribution with the mean value $A^{-1}B$ and the dispersion matrix A^{-1}, the value of the integral I_2 is determined by the expression

$$I_2 = E(\Xi^T U \Xi) \exp\left(\frac{1}{2} B^T A^{-1} B\right), \tag{26}$$

where $E(\Xi^T U \Xi)$ is the mathematical expectation of the quadratic form $\Xi^T U \Xi$ of the random vector Ξ with the Gaussian distribution (23). It is known the equality [2]:

$$\Xi^T U \Xi = \mathrm{Tr}(U \Xi \Xi^T). \tag{27}$$

Then

$$E(\Xi^T U \Xi) = E\left(\mathrm{Tr}(U \Xi \Xi^T)\right) = \mathrm{Tr}\left(U E(\Xi \Xi^T)\right).$$

Further, since for the Gaussian distribution (23) $E(\Xi \Xi^T) = A^{-1} + B^T A^{-1} A^{-1} B,$ then

$$E(\Xi^T U \Xi) = \mathrm{Tr}\left(U(A^{-1} + B^T A^{-1} A^{-1} B)\right) = \mathrm{Tr}(U A^{-1}) + \mathrm{Tr}(U B^T A^{-1} A^{-1} B).$$

If one takes into account the equality of type (27), which can be rewritten as

$$\mathrm{Tr}(U B^T A^{-1} A^{-1} B) = B^T A^{-1} U A^{-1} B,$$

then one receives

$$E(\Xi^T U \Xi) = \mathrm{Tr}(U A^{-1}) + B^T A^{-1} U A^{-1} B,$$

and for the integral I_2:

$$I_2 = \frac{1}{\sqrt{(2\pi)^k |A^{-1}|}} \int\limits_{E^k} \xi^T U \xi \exp\left(-\frac{1}{2} \xi^T A \xi + B^T \xi\right) d\xi = \left(\mathrm{Tr}(U A^{-1}) + B^T A^{-1} U A^{-1} B\right) \exp\left(\frac{1}{2} B^T A^{-1} B\right).$$

As a result, we have received equality (4).

Let us note that equality (4) is more general then the according equality of work [3], since the matrix U in work [3] is supposed to be diagonal.

2. The total probability formula for vector Gaussian distributions.

T h e o r e m 1 (The total probability formula for vector Gaussian distributions). *Let* $\Xi^T = (\Xi_1, \Xi_2, ..., \Xi_k)$ *be a row random vector with k components,* $X^T = (X_1, X_2, ..., X_n)$ *be a row random vector with n components, f(ξ) be the probability density of the vector* Ξ, $f(x/\xi)$ *be the condition probability density of the vector X, and* E^k *be the k-dimensional Euclidean space. If in the total probability formula*

$$f(x) = \int\limits_{E^k} f(x/\xi) f(\xi) d\xi \tag{28}$$

the probability density $f(x/\xi)$ *is represented in the form*

$$f(x/\xi) = \frac{1}{\sqrt{(2\pi)^n |d_X|}} \exp\left(-\frac{1}{2} \xi^T S \xi + V^T \xi - \frac{1}{2} W\right),$$

and the probability density f(ξ) is represented in the form

$$f(\xi) = \frac{1}{\sqrt{(2\pi)^k \, |d_\Xi|}} \exp\left(-\frac{1}{2} \xi^T d_\Xi^{-1} \xi + \nu_\Xi^T d_\Xi^{-1} \xi - \frac{1}{2} \nu_\Xi^T d_\Xi^{-1} \nu_\Xi \right),$$

then integral (28) (*the total probability formula*) *is defined by the following expression*

$$f(x) = \int_{E^k} f(x/\xi) f(\xi) d\xi = \frac{1}{\sqrt{(2\pi)^n \, |d_\Xi A d_X|}} \exp\left(\frac{1}{2} B^T A^{-1} B - \frac{1}{2} C \right), \tag{29}$$

where

$$A = d_\Xi^{-1} + S, \tag{30}$$

$$B = d_\Xi^{-1} \nu_\Xi + V, \tag{31}$$

$$C = \nu_\Xi^T d_\Xi^{-1} \nu_\Xi + W. \tag{32}$$

P r o o f. Performing the multiplication under the integral in (28), we receive

$$f(x/\xi) f(\xi) = \frac{1}{\sqrt{(2\pi)^{n+k} \, |d_X \| d_\Xi|}} \exp\left(-\frac{1}{2} \xi^T A \xi + B^T \xi - \frac{1}{2} C \right), \tag{33}$$

where *A*, *B*, and *C* are defined by formulas (30), (31), and (32). Integration of expression (33) using equality (2) gives expression (29). This completes the proof of theorem 1.

3. The Bayes formula for vector Gaussian distributions.

T h e o r e m 2 (the Bayes formula for vector Gaussian distributions). *Let* $\Xi^T = (\Xi_1, \Xi_2, ..., \Xi_k)$ *be a random vector-row with k components,* $X^T = (X_1, X_2, ..., X_n)$ *be a random vector-row with n components, f(ξ) be the probability density of the vector* Ξ, $f(x/\xi)$ *be the condition probability density of the vector X, and* E^k *be the k-dimensional Euclidean space. If in the Bayes formula*

$$f(\xi/x) = \frac{f(\xi) f(x/\xi)}{\int_{E^k} f(\xi) f(x/\xi) d\xi} \tag{34}$$

the probability density $f(x/\xi)$ *is represented in the form*

$$f(x/\xi) = \frac{1}{\sqrt{(2\pi)^n \, |d_X|}} \exp\left(-\frac{1}{2} \xi^T S \xi + V^T \xi - \frac{1}{2} W \right),$$

and the probability density f(ξ) is represented in the form

$$f(\xi) = \frac{1}{\sqrt{(2\pi)^k \, |d_\Xi|}} \exp\left(-\frac{1}{2} \xi^T d_\Xi^{-1} \xi + \nu_\Xi^T d_\Xi^{-1} \xi - \frac{1}{2} \nu_\Xi^T d_\Xi^{-1} \nu_\Xi \right),$$

then the posteriori probability density $f(\xi/x)$ *of the random vector ξ defined by the Bayes formula* (34) *has the following form*

$$f(\xi/x) = \frac{1}{\sqrt{(2\pi)^k \, |A^{-1}|}} \exp\left(-\frac{1}{2} (\xi - A^{-1} B)^T A (\xi - A^{-1} B) \right), \tag{35}$$

where $A = d_\Xi^{-1} + S$, $B = d_\Xi^{-1} \nu_\Xi + V$.

P r o o f. We note that this theorem is formulated under the same conditions and designations as theorem 1. In this case, the numerator of the Bayes formula (34) is defined by expressions (33), and the denominator is defined by formula (29). Dividing (33) on (29) we receive the formula

$$f(\xi / x) = \frac{1}{\sqrt{(2\pi)^k |A^{-1}|}} \exp\left(-\frac{1}{2}\xi^T A\xi + B^T\xi - \frac{1}{2}B^T A^{-1}B\right), \tag{36}$$

which can be written as in (35). The equality of expressions (36) and (35) is easily verified by multiplying in expression (35). This completes the proof of theorem 2.

Obviously, the expression $A^{-1}B$ in the Bayes formula (34) is the posteriori mathematical expectation of the random vector Ξ, i. e. $A^{-1}B = E(\Xi / x)$, and the matrix A^{-1} is the posteriori dispersion matrix of the random vector Ξ, i. e. $A^{-1} = E\left((\Xi - A^{-1}B)(\Xi - A^{-1}B)^T / x\right)$.

4. Example. As an example, we consider the problem of the calculation of the Bayesian estimators of the coefficients of the multiple regression function.

Let U and Y be the input and output vectors of the controlled object, respectively, and the object is described by the conditional probability density $f(\overline{y} / \overline{\theta}, \overline{u})$. As a rule, it is the Gaussian (normal) probability density:

$$f(\overline{y} / \overline{\theta}, \overline{u}) \sim N\left(\phi(\overline{\theta}, \overline{u}), d_Y\right), \tag{37}$$

where $\phi(\overline{\theta}, \overline{u}) = \overline{y}$ is the regression function of Y on U, $\overline{u}$ and $\overline{y}$ are the input and output vectors of the regression function, respectively, $\overline{\theta}$ is the vector of the coefficients of the regression function, and d_Y is the constant dispersion matrix of the internal noise of the object. Description (37) could be represented in the form

$$Y = \phi(\overline{\theta}, \overline{u}) + E,$$

where E is the random vector with the Gaussian distribution $N(0, d_E)$.

The multiple regression function is considered most frequently when $\overline{y}$ is scalar (we will denote it y) and $\overline{u}$ is vector.

The class of the functions is represented in the form

$$y = \phi(\overline{\theta}, \overline{u}) = \sum_{j=1}^{m} h_j(\overline{u})\theta_j = \overline{h}^T\overline{\theta},$$

where $h_j(\overline{u})$, $j = 1,2,...,m$, are basis functions, $\overline{h}^T = \overline{h}^T(\overline{u}) = \left(h_1(\overline{u}), h_2(\overline{u}),..., h_m(\overline{u})\right)$ is the vector formed by the basis functions, and $\overline{\theta}^T = (\theta_1, \theta_2,..., \theta_m)$ is the vector of the coefficients of the regression function, is usually used to describe the multiple regression function. For example, if we want to write in the vector form a function of two variables u_1, u_2 having the form

$$y = \alpha + \beta u_1 + \gamma u_2 + \tau u_1^2,$$

then we have to choose $\overline{h}^T = (1, u_1, u_2, u_1^2)$, $\overline{\theta}^T = (\alpha, \beta, \gamma, \tau)$.

Let $y_{o,i} = \overline{h}_i^T \Theta + \varepsilon_i$, $\Theta^T = (\Theta_1, \Theta_2,..., \Theta_m)$, be the i-th observed value of the output variable Y of the object on the observed value $\overline{h}_i^T$ of the input vector of the basis functions, $i = 1,2,...,n$, and the vector Θ has the normal priory probability density $N(a_\Theta, d_\Theta)$. The task is to find the estimation $\widehat{\theta}$ of the vector $\overline{\theta}$ on the basis of the observers $(\overline{h}_1, y_{o,1})$, $(\overline{h}_2, y_{o,2}),...,(\overline{h}_n, y_{o,n})$ provided $\Theta = \overline{\theta}$.

In our case we have the following probability density functions:

$$f(\overline{\theta}) = \frac{1}{\sqrt{(2\pi)^m d_\Theta}} \exp\left(-\frac{1}{2}(\overline{\theta} - a_\Theta)^T d_\Theta^{-1}(\overline{\theta} - a_\Theta)\right),$$

$$f(y / \overline{\theta}) = \frac{1}{\sqrt{2\pi d_{\mathrm{E}}}} \exp\left(-\frac{1}{2}(y - \overline{h}^{\,T}\overline{\theta})^T d_{\mathrm{E}}^{-1}(y - \overline{h}^{\,T}\overline{\theta}) \right).$$

The vector of observations $\vec{y} = (y_{o,1}, y_{o,2}, ..., y_{o,n})$ will have the following probability density function:

$$f(\vec{y}_o / \overline{\theta}) = \prod_{i=1}^{n} f(y_{o,i} / \overline{\theta}) \sim \frac{1}{\sqrt{(2\pi)^n d_{\mathrm{E}}^n}} \exp\left(-\sum_{i=1}^{n} \frac{1}{2}(y_{o,i} - \overline{h}_i^{\,T}\overline{\theta})^T d_{\mathrm{E}}^{-1}(y_{o,i} - \overline{h}_i^{\,T}\overline{\theta}) \right).$$

We find now the posterior probability density functions $f(\overline{\theta} / \vec{y}_o)$ of the vector coefficient $\overline{\theta}$ by the Bayes formula:

$$f(\overline{\theta} / \vec{y}_o) = \frac{f(\overline{\theta}) f(\vec{y}_o / \overline{\theta})}{\int\limits_{E^m} f(\overline{\theta}) f(\vec{y}_o / \overline{\theta}) d\overline{\theta}}.$$

We will use for this theorem 2. Since

$$f(\overline{\theta}) = \frac{1}{\sqrt{(2\pi)^m d_\Theta}} \exp\left(-\frac{1}{2}\overline{\theta}^T d_\Theta^{-1}\overline{\theta} + a_\Theta^T d_\Theta^{-1}\overline{\theta} - \frac{1}{2} a_\Theta^T d_\Theta^{-1} a_\Theta \right),$$

$$f(\vec{y}_o / \overline{\theta}) = \frac{1}{\sqrt{(2\pi)^n d_{\mathrm{E}}^n}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} y_{o,i}^T d_{\mathrm{E}}^{-1} y_{o,i} + \sum_{i=1}^{n} y_{o,i}^T d_{\mathrm{E}}^{-1}\overline{h}_i^{\,T}\overline{\theta} - \frac{1}{2}\sum_{i=1}^{n} \overline{\theta}^T \overline{h}_i d_{\mathrm{E}}^{-1}\overline{h}_i^{\,T}\overline{\theta} \right),$$

then, in accordance with theorem 2, we have

$$f(\overline{\theta} / \vec{y}_o) = \frac{1}{\sqrt{(2\pi)^m |A^{-1}|}} \exp\left(-\frac{1}{2}(\overline{\theta} - A^{-1}B)^T A(\overline{\theta} - A^{-1}B) \right),$$

where

$$A = d_\Theta^{-1} + \sum_{i=1}^{n} \overline{h}_i d_{\mathrm{E}}^{-1}\overline{h}_i^{\,T}, \quad B = d_\Theta^{-1} a_\Theta + \sum_{i=1}^{n} h_i d_{\mathrm{E}}^{-1} y_{o,i}.$$

Provided the loss function is quadratic, $W(\widehat{\overline{\theta}}, \Theta) = (\widehat{\overline{\theta}} - \Theta)^T (\widehat{\overline{\theta}} - \Theta)$, we get the Bayesian estimation $\widehat{\overline{\theta}}$ of the vector $\overline{\theta}$: $\widehat{\overline{\theta}} = A^{-1}B$.

Conclusion. The results represented in this article provide a basis for theoretical solutions of the vector problems formulated within the framework of the statistical decision theory. The integrals can also be used as table integrals. The possible generalizations of the obtained results for solving more complicated problems within the framework of the statistical decision theory are of great interest.

References

1. Fel'dbaum A. A. *Optimal Control Systems.* New York, London, Academic Press, 1965. 452 p.
2. Rao C. S. *Linear Statistical Inference and its Applications*, 2nd ed., Wiley, 1973. 648 p.
3. Mukha V. S. Calculation of integrals connected with the multivariate Gaussian distribution. *Proceedings of the LETI*, 1974, vol. 160, pp. 27–30 (in Russian).
4. Gantmacher F. R. *The Theory of Matrices.* New York, Chelsea Publishing Company, 1959. Vol. 1. 374 p.
5. Rudin W. *Principles of Mathematical Analysis*, 3ed ed., New York, McGraw-Hill Inc., 1976. 352 p.
6. Prudnikov A. P., Brychkov Yu. A., Marichev O. I. *Integrals and Series.* Translated from the Russian by N. M. Queen. New York, Gordon and Breach Science Publ., 1986. 753 p.

Information about the authors

Vladimir S. Mukha – Dr. Sc. (Engineering), Professor, Professor of the Department of Information Technologies of Automated Systems, Belarusian State University of Informatics and Radioelectronics (6, P. Brovka Str., 220013, Minsk, Republic of Belarus). E-mail: mukha@bsuir.by

Nancy Farat Kako – Postgraduate Student, Belarusian State University of Informatics and Radioelectronics (6, P. Brovka Str., 220013, Minsk, Republic of Belarus). E-mail: kako.nancy@gmail.com

Информация об авторах

Муха Владимир Степанович – доктор технических наук, профессор, профессор кафедры информационных технологий автоматизированных систем, Белорусский государственный университет информатики и радиоэлектроники (ул. П. Бровки, 6, 220013, г. Минск, Республика Беларусь). E-mail: mukha@bsuir.by

Како Нэнси Фарат – аспирант, Белорусский государственный университет информатики и радиоэлектроники (ул. П. Бровки, 6, 220013, г. Минск, Республика Беларусь). E-mail: kako.nancy@gmail.com