Ann. Inst. Statist. Math.
Vol. 55, No. 3, 507–522 (2003)

ON MULTIVARIATE GAUSSIAN TAILS

ENKELEJD HASHORVA AND JÜRG HÜSLER

Department of Mathematical Statistics and Actuarial Sciences, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland, e-mail: enkelejd.hashorva@Allianz-Suisse.ch; juerg.huesler@stat.unibe.ch

(Received February 15, 2001; revised October 7, 2002)

Abstract. Let $\{\boldsymbol{X}_n, n \geq 1\}$ be a sequence of standard Gaussian random vectors in $I\!R^d$, $d \geq 2$. In this paper we derive lower and upper bounds for the tail probability $\boldsymbol{P}\{\boldsymbol{X}_n > \boldsymbol{t}_n\}$ with $\boldsymbol{t}_n \in I\!R^d$ some threshold. We improve for instance bounds on Mills ratio obtained by Savage (1962, *J. Res. Nat. Bur. Standards Sect. B*, **66**, 93–96). Furthermore, we prove exact asymptotics under fairly general conditions on both $\boldsymbol{X}_n$ and $\boldsymbol{t}_n$, as $\|\boldsymbol{t}_n\| \to \infty$ where the correlation matrix Σ_n of $\boldsymbol{X}_n$ may also depend on n.

Key words and phrases: Multivariate Mills ratio, Gaussian random sequences, tail asymptotics, quadratic programming.

1. Introduction

Let $\boldsymbol{X}$ be a Gaussian random vector in $I\!R^d, d \geq 1$ with underlying covariance matrix Σ and $\boldsymbol{t} \in I\!R^d$ some threshold. Multivariate Mills ratio (see e.g. Tong (1989)) is defined by

$$R(\boldsymbol{t}, \Sigma) := \boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}\}/\varphi(\boldsymbol{t}) = \int_{(0,\infty)^d} \exp(-\langle \boldsymbol{x} + 2\boldsymbol{t}, \Sigma^{-1}\boldsymbol{x}\rangle/2) d\boldsymbol{x}, \tag{1.1}$$

with $\varphi(\boldsymbol{t}) := \exp(-\langle \boldsymbol{t}, \Sigma^{-1}\boldsymbol{t}\rangle/2)(2\pi)^{-d/2}|\Sigma|^{-1/2}$ the density function of $\boldsymbol{X}$ and $\langle\cdot,\cdot\rangle$ the scalar product in $I\!R^d$. In the univariate case the Laplace-Feller inequality reads (cf. Barndorff-Nielsen and Cox (1989), p. 56)

$$\begin{aligned} R(t) := R(t, 1) &= \int_0^\infty \exp\left(-\frac{s^2}{2} - st\right) ds \\ &\doteq \frac{1}{t}\left[1 - \frac{1}{t^2} + \frac{1 \cdot 3}{t^4} + \cdots\right], \qquad t > 0, \end{aligned} \tag{1.2}$$

whereas from Stanislaw and Werner (1999), the following inequality holds for $t > -1$

$$\frac{2}{t + (t^2+4)^{1/2}} \leq R(t) \leq \frac{4}{3t + (t^2+8)^{1/2}}. \tag{1.3}$$

In the multivariate setup, the covariance matrix Σ plays a crucial role. Clearly, we may consider without loss of generality (w.l.o.g.) standard Gaussian random vectors, i.e. the components of the random vector $\boldsymbol{X}$ are normally distributed with mean 0 and unit standard deviation.

By Savage (1962) (see also Tong (1989)) if

$$\Sigma^{-1}\boldsymbol{t} > \boldsymbol{0} \tag{1.4}$$

holds, then

$$\frac{1 - \langle 1/(\Sigma^{-1}\boldsymbol{t}), \Sigma^{-1}(1/(\Sigma^{-1}\boldsymbol{t}))\rangle}{\prod_{i=1}^{d} h_i} \le R(\boldsymbol{t}, \Sigma) \le \frac{1}{\prod_{i=1}^{d} h_i}, \qquad \boldsymbol{t} \in \mathbb{R}^d, d \ge 2, \tag{1.5}$$

with $\boldsymbol{0} := (0, \ldots, 0)' \in \mathbb{R}^d$, $h_i := \langle \boldsymbol{e}_{ii}, \Sigma^{-1}\boldsymbol{t}\rangle > 0$, $i = 1, \ldots, d$ and $\boldsymbol{e}_i$ the i-th unit vector in $\mathbb{R}^d$.

Knowing the Gaussian density function, upper and lower bounds for Mills ratio are easily converted to bounds for the tail probability $\boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}\}$.

One of the merit of the above inequality is that when (1.4) holds and $h_i \to \infty$ for all $i = 1, \ldots, d$ then we obtain the exact rate of convergence to 0 of $\boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}\}$. This is the case for example if Σ^{-1} has all entries positive and $\boldsymbol{t} \in (0, \infty)^d$.

Under further restrictions, assuming that $t_i \to \infty$ for some i and the matrix Σ^{-1} has positive entries

$$R(\boldsymbol{t}, \Sigma) > \frac{1}{\prod_{i=1}^{d} h_i} - \frac{1}{\prod_{i=1}^{d} h_i^{1+1/d}}, \qquad \boldsymbol{t} \in (0, \infty)^d \tag{1.6}$$

is obtained by Gjacjauskas (1973).

As pointed out by Steck (1979) condition (1.4) is rather restrictive. Several upper and lower bounds are derived in Steck (1979) and Satish (1986). They are easy to calculate numerically, however they do not give an explicit idea what happens for large thresholds in the sense of the speed of convergence to 0 for the probability of interest. Fang and Xu (1990) consider generalisation of Mills ratio for spherically symmetric distribution functions.

Of course there exist numerous approaches including simulations to calculate the probability of observing a Gaussian vector above a given boundary. However, if one is interested in the speed of convergence to 0 letting $\langle \boldsymbol{t}, \boldsymbol{t}\rangle$ tend to ∞, such methods do not provide the answer.

The main purpose of this paper is to establish new upper and lower bounds for $\boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}\}$, (hence for the multivariate Mills ratio) that inherit the simplicity of (1.5) and hold in general even if condition (1.4) is not satisfied. In a second attempt we treat then the asymptotic behavior of $\boldsymbol{P}\{\boldsymbol{X}_n \ge \boldsymbol{t}_n\}$ (for $n \to \infty$) where $\boldsymbol{X}_n$ has correlation matrix Σ_n varying with n. The upper bound given in the three dimensional case by Proposition 4.2 of Raab (1999), Lemma D in Shao and Santosh (1999) as well as part of Lemma 2.1 of Elnaggar and Mukherjea (1999) can be obtained as special cases of our results.

2. Notations and preliminaries

Let in the sequel $I \subset \{1, \ldots, d\}$ with $d \ge 2$ denotes some non-empty index set. Further put $J := \{1, \ldots, d\} \setminus I$. The number of elements for an index sets I is denoted by $|I| := \text{card}\{I\}$. The symbol Σ will be reserved for positive definite correlation matrices. Clearly its inverse matrix exists, denoted by B throughout.

Abusing slightly the notations we drop the transpose sign when writing vectors, and also drop the subscripts for vectors in $\mathbb{R}^d$. For a given vector $\boldsymbol{x} \in \mathbb{R}^d$ we write $\boldsymbol{x}_I$ the vector obtained by deleting the components of $\boldsymbol{x}$ in J. Similar notations A_{II}, A_{IJ}, A_{JI}, A_{JJ} are used for submatrices of a given matrix $A \in I\!R^{d\times d}$. In our notation indexing is performed first so for example the notation A_{II}^{-1} means $(A_{II})^{-1}$. The following standard notations will be used frequently

$$\begin{aligned}
&\mathbf{1} := (1,\dots,1)' \in I\!R^d, \qquad \boldsymbol{\infty} = (\infty,\dots,\infty)' \in I\!R^d,\\
&\boldsymbol{x} > \boldsymbol{y}, \quad \text{if} \quad x_i > y_i, \qquad \forall i = 1,\dots,d,\\
&\boldsymbol{x} \ge \boldsymbol{y}, \quad \text{if} \quad x_i \ge y_i, \qquad \forall i = 1,\dots,d,\\
&|\boldsymbol{x}|^2 := \langle \boldsymbol{x}, \boldsymbol{x}\rangle, \qquad \boldsymbol{x} \in I\!R^d,\\
&c\boldsymbol{x} := (cx_1,\dots,cx_d)', \qquad c \in I\!R,\\
&\boldsymbol{a}\boldsymbol{x} := \operatorname{diag}(\boldsymbol{a})\boldsymbol{x} = (a_1x_1,\dots,a_dx_d)', \qquad \boldsymbol{a} \in I\!R^d,
\end{aligned}$$

with $\operatorname{diag}(\boldsymbol{a})$ the diagonal matrix corresponding to $\boldsymbol{a}$.

To this end, we provide a minor generalisation of Proposition 2.5 of Hashorva and Hüsler (2002) which is a crucial tool for the rest of the paper. Its proof is omitted since it follows easily by slightly modifying the proof of the aforementioned proposition.

PROPOSITION 2.1. *Let $\Sigma \in I\!R^{d\times d}$ be a positive definite correlation matrix and $\boldsymbol{x}^*$ the unique solution of the quadratic programming $(\mathcal{P}_{B,\boldsymbol{t}})$: minimise $\langle \boldsymbol{x}, B\boldsymbol{x}\rangle$ under $\boldsymbol{x} \ge \boldsymbol{t}$ with $\boldsymbol{t} \notin (-\infty, 0]^d$. Then there exists a unique index set $I_t \subset \{1,\dots,d\}$ so that*

(2.1) $$1 \le |I_t| \le d$$

(2.2) $$\boldsymbol{x}^*_{I_t} = \boldsymbol{t}_{I_t} \ne \mathbf{0}_{I_t}, \quad \text{and if} \quad J_t \ne \emptyset \quad \boldsymbol{x}^*_{J_t} = -B^{-1}_{J_tJ_t}B_{J_tJ_t}\boldsymbol{t}_{I_t} \ge \boldsymbol{t}_{J_t},$$

(2.3) $$\forall i \in I \quad h_i := \langle \boldsymbol{e}_i, \Sigma^{-1}_{I_tI_t}\boldsymbol{t}_{I_t}\rangle > 0$$

(2.4) $$\alpha_t = \min_{\boldsymbol{x}\ge\boldsymbol{t}} \langle \boldsymbol{x}, B\boldsymbol{x}\rangle = \langle \boldsymbol{x}^*, \Sigma^{-1}\boldsymbol{x}^*\rangle = \langle \boldsymbol{t}_{I_t}, \Sigma^{-1}_{I_tI_t}\boldsymbol{t}_{I_t}\rangle > 0,$$

with $J_t := \{1,\dots,d\} \setminus I_t$, and $\boldsymbol{e}_i$ the i-th unit vector in $I\!R^{|I|}$.

PROOF. The claim follows along the lines of the proof of Proposition 2.5 in Hashorva and Hüsler (2002), therefore omitted here. □

The role of the index set I_t is very crucial for our discussion. In terms of geometry, $|I_t|$ is the codimension of the intersection of the tangent hyperplane at $\boldsymbol{x}^*$ to the ellipsoid $\alpha = \langle \boldsymbol{x}, B\boldsymbol{x}\rangle$ with the boundary of the set $[\boldsymbol{t}, \boldsymbol{\infty})$. For notation simplicity we drop the subscript for all $\boldsymbol{t}$ and write simply I, J instead of I_t, J_t. It is very important to remember that the index set varies with $\boldsymbol{t}$, and so does the unique solution of quadratic problem $\boldsymbol{x}^*$. In our discussion, the following index set is also important

$$J^* := \{j \in J : \boldsymbol{x}^*_j = \boldsymbol{t}_j\}.$$

From (2.2) we know that $\boldsymbol{x}^*_J \ge \boldsymbol{t}_J$ so it may happen that J^* is empty. Clearly this is the case when J is empty. The extreme case $J^* = J$ is also possible; it is shown in Example 1 in the next section. If $\boldsymbol{t} = t\mathbf{1}$ with t some positive constant we have by Proposition 2.5 of Hashorva and Hüsler (2002)

(2.5) $$2 \le |I| \le d,$$

hence for the bivariate case ($d = 2$) we simply have $J = J^* = \emptyset$ and $\boldsymbol{x}^* = t(1,1)$ is the unique solution of the quadratic programming problem $(\mathcal{P}_{B,t(1,1)})$.

3. Upper and lower bounds

As pointed out in the introduction, condition (1.4) is crucial to derive the Savage upper and lower bounds. It is of interest to treat also the less common case when the mentioned condition does not hold. Since the upper Savage bound becomes too large for $|\boldsymbol{t}|$ close to 0, on the other side, it is asymptotically exact if $\boldsymbol{t} \to \infty$, we construct an upper bound which is accurate for both cases.

Let $\boldsymbol{X}$ be some standard Gaussian random vector on $\mathbb{R}^d$, $d \geq 2$, with underlying correlation matrix Σ. Define according to Proposition 2.1 α_t, I, J, h_i, Σ_{II}, $\boldsymbol{x}^*$ with respect to some threshold $\boldsymbol{t} \notin (-\infty, 0]^d$. Next, we have the obvious inequality

$$\boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}\} \leq \boldsymbol{P}\{\boldsymbol{X}_I > \boldsymbol{t}_I\},$$

hence in light of (1.5)

$$\boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}\} \leq \frac{\exp(-\alpha_t/2)}{(2\pi)^{|I|/2}|\Sigma_{II}|^{1/2}} \prod_{i \in I} h_i^{-1}. \tag{3.1}$$

Note in passing that the right hand side above does not depend on $\boldsymbol{t}_J$.

The upper bound in (3.1) seems to be rather crude. Actually, if $J^* = \emptyset$ it is exact asymptotically (for $h_i \to \infty$) as will be shown in the next section. For $\boldsymbol{t}$ so that h_i are close to 0, clearly this upper bound is far beyond 1, therefore of no use. This case will be dealt with, too.

In the next proposition we derive another upper bound of the same asymptotic order which depends on $\boldsymbol{t}_J$. Further, the case $|\boldsymbol{t}|$ close to 0 is also covered. To this end we put $\boldsymbol{P}\{\boldsymbol{X}_J > (\boldsymbol{t} - \boldsymbol{x}^*)_J\}|\Sigma_{JJ}|^{1/2} =: 1$ and $\prod_{i \in J}(\cdots) =: 1$ whenever $J = \emptyset$.

PROPOSITION 3.1. *Under the previous assumptions and notations we have*

$$\boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}\} \leq \frac{\exp(-\alpha_t/2)}{(2\pi)^{|I|/2}|\Sigma|^{1/2}} \min(c_0(\boldsymbol{t}), c_1(\boldsymbol{t}), c_2(\boldsymbol{t})), \tag{3.2}$$

where

$$c_0(\boldsymbol{t}) := \prod_{i \in I} h_i^{-1} \sqrt{|\Sigma|/|\Sigma_{II}|},$$

$$c_1(\boldsymbol{t}) := \prod_{i \in I} R(h_i \sqrt{\lambda_\vee}) \lambda_\vee^{d/2} \prod_{i \in J} \boldsymbol{P}\{X_i > (t_i - x_i^*)/\sqrt{\lambda_\vee}\},$$

$$c_2(\boldsymbol{t}) := \prod_{i \in I} h_i^{-1} \boldsymbol{P}\{\boldsymbol{X}_J > (\boldsymbol{t} - \boldsymbol{x}^*)_J\}|\Sigma_{JJ}|^{1/2},$$

with $\lambda_\vee > 0$ *the maximal eigenvalue of* Σ *and* $\boldsymbol{t} - \boldsymbol{x}^* \leq \boldsymbol{0}$.

PROOF. The first bound with $c_0(\boldsymbol{t})$ follows immediately from (3.1).

For sake of simplicity we assume that $J \neq \emptyset$; the case $J = \emptyset$ is simpler and its proof is therefore omitted. Now recall that both square matrices B_{II}, B_{JJ} are positive definite and moreover

$$\Sigma_{II}^{-1} = B_{II} - B_{IJ} B_{JJ}^{-1} B_{JI}, \qquad \Sigma_{JJ}^{-1} = B_{JJ} - B_{JI} B_{II}^{-1} B_{IJ}. \tag{3.3}$$

By Proposition 2.1 we have $\boldsymbol{x}_I^* = \boldsymbol{t}_I, \boldsymbol{x}_J^* = -B_{JJ}^{-1}B_{JJ}\boldsymbol{t}_I \geq \boldsymbol{t}_J$, hence we obtain

$$\begin{aligned}
\langle \boldsymbol{x}+\boldsymbol{x}^*, B(\boldsymbol{x}+\boldsymbol{x}^*)\rangle &= \langle \boldsymbol{x}_J + \boldsymbol{x}_J^* + B_{JJ}^{-1}B_{JI}(\boldsymbol{x}_I+\boldsymbol{t}_I), \\
&\quad B_{JJ}(\boldsymbol{x}_J + \boldsymbol{x}_J^* + B_{JJ}^{-1}B_{JI}(\boldsymbol{x}_I+\boldsymbol{t}_I))\rangle \\
&\quad + \langle \boldsymbol{x}_I+\boldsymbol{t}_I, (B_{II} - B_{IJ}B_{JJ}^{-1}B_{JI})(\boldsymbol{x}_I+\boldsymbol{t}_I)\rangle \\
&= \langle \boldsymbol{x}_J + (\boldsymbol{x}_J^* + B_{JJ}^{-1}B_{JI}\boldsymbol{t}_I) + B_{JJ}^{-1}B_{JI}\boldsymbol{x}_I, \\
&\quad B_{JJ}(\boldsymbol{x}_J + (\boldsymbol{x}_J^* + B_{JJ}^{-1}B_{JI}\boldsymbol{t}_I) + B_{JJ}^{-1}B_{JI}\boldsymbol{x}_I)\rangle \\
&\quad + \langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{x}_I\rangle + 2\langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle + \langle \boldsymbol{t}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle \\
&= \langle \boldsymbol{x}_J + B_{JJ}^{-1}B_{JI}\boldsymbol{x}_I, B_{JJ}(\boldsymbol{x}_J + B_{JJ}^{-1}B_{JI}\boldsymbol{x}_I)\rangle + \langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{x}_I\rangle \\
&\quad + 2\langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle + \alpha_t \\
&= \langle \boldsymbol{x}, B\boldsymbol{x}\rangle + 2\langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle + \alpha_t. \qquad (3.4)
\end{aligned}$$

Since the correlation matrix Σ is positive definite, there exists an orthogonal matrix D such that $\Sigma = D\Lambda D$ with $\Lambda := \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_d)$ is the diagonal matrix of the positive eigenvalues of Σ. So we may write

$$B = D\Lambda^{-1}D = D\operatorname{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \ldots, \lambda_d^{-1})D.$$

Let $\lambda_\vee$, $\lambda_\wedge$ be the largest and the smallest eigenvalue of Σ respectively. Clearly $d \geq \lambda_\vee \geq \lambda_\wedge > 0$ and moreover for all $\boldsymbol{x} \in \mathbb{R}^d$

$$\frac{1}{\lambda_\vee}\langle \boldsymbol{x}, \boldsymbol{x}\rangle \leq \langle \boldsymbol{x}, B\boldsymbol{x}\rangle \leq \frac{1}{\lambda_\wedge}\langle \boldsymbol{x}, \boldsymbol{x}\rangle, \qquad (3.5)$$

hence (3.4) implies

$$\langle \boldsymbol{x}+\boldsymbol{x}^*, B(\boldsymbol{x}+\boldsymbol{x}^*)\rangle \geq \frac{1}{\lambda_\vee}\langle \boldsymbol{x}, \boldsymbol{x}\rangle + 2\langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle + \alpha_t.$$

In light of Proposition 2.1

$$\begin{aligned}
&(\boldsymbol{t}-\boldsymbol{x}^*)_J \leq \boldsymbol{0}_J, \qquad (\boldsymbol{t}-\boldsymbol{x}^*)_I = \boldsymbol{0}_I \\
&h_i := \langle \boldsymbol{e}_i, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle > 0, \qquad \forall i \in I,
\end{aligned}$$

with $\boldsymbol{e}_i$ the unit vector in $\mathbb{R}^{|I|}$, hence transforming the variables yields using (3.4) and (3.5)

$$\begin{aligned}
\boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}\} &= \int_{\boldsymbol{x}>\boldsymbol{t}} \varphi(\boldsymbol{x})d\boldsymbol{x} \\
&= \int_{\boldsymbol{x}>\boldsymbol{t}-\boldsymbol{x}^*} \varphi(\boldsymbol{x}+\boldsymbol{x}^*)d\boldsymbol{x} \\
&\leq \frac{\exp(-\alpha_t/2)}{(2\pi)^{d/2}|\Sigma|^{1/2}} \int_{\boldsymbol{x}_J>(\boldsymbol{t}-\boldsymbol{x}^*)_J} \int_{(0,\infty)^{|I|}} \exp(-\langle \boldsymbol{x}, \boldsymbol{x}\rangle/(2\lambda_\vee) - \langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle)d\boldsymbol{x} \\
&= \frac{\exp(-\alpha_t/2)}{(2\pi)^{d/2}|\Sigma|^{1/2}} \int_{\boldsymbol{x}_J>(\boldsymbol{t}-\boldsymbol{x}^*)_J} \exp(-\langle \boldsymbol{x}_J, \boldsymbol{x}_J\rangle/(2\lambda_\vee))d\boldsymbol{x}_J \\
&\quad \times \int_{(0,\infty)^{|I|}} \exp(-\langle \boldsymbol{x}_I, \boldsymbol{x}_I\rangle/(2\lambda_\vee) - \langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle)d\boldsymbol{x}_I
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\exp(-\alpha_t/2)\lambda_{\vee}^{d/2}}{(2\pi)^{d/2}|\Sigma|^{1/2}} \int_{\boldsymbol{x}_J > (\boldsymbol{t}-\boldsymbol{x}^*)_J/\sqrt{\lambda_\vee}} \exp(-\langle \boldsymbol{x}_J, \boldsymbol{x}_J\rangle/2) d\boldsymbol{x}_J \\
&\quad \times \prod_{i\in I} \int_0^\infty \exp(-s^2/2 - \sqrt{\lambda_\vee} h_i s) ds \\
&= \frac{\exp(-\alpha_t/2)\prod_{i\in I} R(\sqrt{\lambda_\vee} h_i)}{(2\pi)^{d/2}|\Sigma|^{1/2}} \left(\lambda_\vee^{d/2} \prod_{i\in J} \boldsymbol{P}\{X_i > (t_i - x_i^*)/\sqrt{\lambda_\vee}\}\right).
\end{aligned}
$$

Next, using now (3.3) and the fact that the square matrix B_{II} is positive definite

$$
\begin{aligned}
(3.4) &= \langle \boldsymbol{x}_J, (B_{JJ} - B_{JI}B_{II}^{-1}B_{IJ})\boldsymbol{x}_J\rangle + \langle \boldsymbol{x}_I + B_{II}^{-1}B_{IJ}\boldsymbol{x}_J, B_{II}(\boldsymbol{x}_I + B_{II}^{-1}B_{IJ}\boldsymbol{x}_J)\rangle \\
&\quad + 2\langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle + \alpha_t \\
&= \langle \boldsymbol{x}_J, \Sigma_{JJ}^{-1}\boldsymbol{x}_J\rangle + \langle \boldsymbol{x}_I + B_{II}^{-1}B_{IJ}\boldsymbol{x}_J, B_{II}(\boldsymbol{x}_I + B_{II}^{-1}B_{IJ}\boldsymbol{x}_J)\rangle + 2\langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle + \alpha_t \\
(3.6) \quad &\geq \langle \boldsymbol{x}_J, \Sigma_{JJ}^{-1}\boldsymbol{x}_J\rangle + 2\langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle + \alpha_t.
\end{aligned}
$$

Proceeding similarly we may write

$$
\begin{aligned}
\boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}\} &\leq \frac{\exp(-\alpha_t/2)}{(2\pi)^{d/2}|\Sigma|^{1/2}} \int_{\boldsymbol{x}_J > (\boldsymbol{t}-\boldsymbol{x}^*)_J} \int_{(0,\infty)^{|I|}} \exp(-\langle \boldsymbol{x}_J, \Sigma_{JJ}^{-1}\boldsymbol{x}_J\rangle/2 - \langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle) d\boldsymbol{x} \\
&= \frac{\exp(-\alpha_t/2)}{(2\pi)^{d/2}|\Sigma|^{1/2}} \int_{\boldsymbol{x}_J > (\boldsymbol{t}-\boldsymbol{x}^*)_J} \exp(-\langle \boldsymbol{x}_J, \Sigma_{JJ}^{-1}\boldsymbol{x}_J\rangle/2) d\boldsymbol{x}_J \prod_{i\in I} \int_0^\infty \exp(-h_i s) ds \\
&= \frac{\exp(-\alpha_t/2)\prod_{i\in I} h_i^{-1}}{(2\pi)^{|I|/2}|\Sigma|^{1/2}|\Sigma_{JJ}|^{-1/2}} \int_{\boldsymbol{x}_J > (\boldsymbol{t}-\boldsymbol{x}^*)_J} \frac{\exp(-\langle \boldsymbol{x}_J, \Sigma_{JJ}^{-1}\boldsymbol{x}_J\rangle/2)}{(2\pi)^{|J|/2}|\Sigma_{JJ}|^{1/2}} d\boldsymbol{x}_J,
\end{aligned}
$$

hence the proof. □

Let us briefly consider the role of $c_0(\boldsymbol{t}), c_1(\boldsymbol{t}), c_2(\boldsymbol{t})$ for the upper bound above. Since

$$
\boldsymbol{P}\{X_i > (t - x_i^*)/\sqrt{\lambda_\vee}\} \geq 1/2, \ \forall i \in J
$$

and recalling that h_i, $i \in I$ are positive we get applying (1.3)

$$
\begin{aligned}
0 &< \lambda_\vee^{d/2} 2^{|I|-|J|} \prod_{i\in I} \frac{1}{h_i\sqrt{\lambda_\vee} + (h_i^2\lambda_\vee + 4)^{1/2}} \leq c_1(\boldsymbol{t}) \\
&\leq \lambda_\vee^{d/2} 4^{|I|} \prod_{i\in I} \frac{1}{3h_i\sqrt{\lambda_\vee} + (h_i^2\lambda_\vee + 8)^{1/2}},
\end{aligned}
$$

and since $\boldsymbol{t} - \boldsymbol{x}^* \leq \boldsymbol{0}$

$$
\boldsymbol{P}\{\boldsymbol{X}_J > \boldsymbol{0}_J\}|\Sigma_{JJ}|^{1/2} \prod_{i\in I} h_i^{-1} \leq c_2(\boldsymbol{t}) \leq |\Sigma_{JJ}|^{1/2} \prod_{i\in I} h_i^{-1}
$$

holds. It is obvious that for h_i close to 0, $c_0(\boldsymbol{t})$, $c_2(\boldsymbol{t})$ become large, whereas $c_1(\boldsymbol{t})$ remains unaffected, implying $\min(c_0(\boldsymbol{t}), c_1(\boldsymbol{t}), c_2(\boldsymbol{t})) = c_1(\boldsymbol{t})$. On the other side, if $h_i \to \infty$ then all three constants $c_0(\boldsymbol{t})$, $c_1(\boldsymbol{t})$, $c_2(\boldsymbol{t})$ capture the speed of convergence to 0 of the tail probability, see discussion in the next section. The merit of both $c_0(\boldsymbol{t})$, $c_2(\boldsymbol{t})$ is that they do not depend on the maximal eigenvalue of Σ.

Let in the following $\bar{h}_i = h_i/\sqrt{b_{ii}}$ and $\bar{B}$ the matrix B with main diagonal entries 0; recall that also the minimal eigenvalue $\lambda_\wedge$ of the correlation matrix Σ is positive. Now we formulate the result concerning the lower bound:

PROPOSITION 3.2. *With the same conditions and notations as in Proposition* 3.1 *we have*

$$P\{X > t\} \geq \frac{\exp(-\alpha_t/2)}{(2\pi)^{|I|/2}|\Sigma|^{1/2}} \max(c_3(t), c_4(t)) \tag{3.7}$$

where

$$\begin{aligned} c_3(t) &:= \prod_{i\in I} R(h_i\sqrt{\lambda_\wedge})\lambda_\vee^{d/2} \prod_{i\in J} P\{X_i > (t_i - x_i^*)/\sqrt{\lambda_\wedge}\} \\ &\geq \prod_{i\in I} R(h_i\sqrt{\lambda_\wedge})\lambda_\vee^{d/2} 2^{-|J|} > 0, \\ c_4(t) &:= \exp(-\langle V, \bar{B}V\rangle/2) \prod_{i\in I} \frac{R(\bar{h}_i)}{\sqrt{b_{ii}}} \prod_{i\in J} \frac{P\{X_i > (t - x_i^*)\sqrt{b_{ii}}\}}{\sqrt{b_{ii}}} \\ &\geq \exp(-\langle V, \bar{B}V\rangle/2) 2^{-|J|} \prod_{i\in I} R(\bar{h}_i) \prod_{i=1,\ldots,d} \frac{1}{\sqrt{b_{ii}}} > 0, \end{aligned}$$

with

$$V_i := \begin{cases} [1/R(\bar{h}_i) - \bar{h}_i] b_{ii}^{-1/2}, & \textit{for} \quad i \in I, \\ R((t_i - x_i^*)\sqrt{b_{ii}})^{-1} b_{ii}^{-1/2}, & \textit{for} \quad i \in J. \end{cases}$$

PROOF. As in the proof of the previous proposition we assume for simplicity again $J \neq \emptyset$. In light of (3.4) and (3.5) we get

$$\langle x + x^*, B(x + x^*)\rangle \leq \frac{1}{\lambda_\wedge}\langle x, x\rangle + 2\langle x_I, \Sigma_{II}^{-1} t_I\rangle + \alpha_t, \qquad x \in \mathbb{R}^d,$$

thus analogously to the proof above

$$P\{X > t\} \geq \frac{\exp(-\alpha_t/2)}{(2\pi)^{|I|/2}|\Sigma|^{1/2}} \prod_{i\in I} R(h_i\sqrt{\lambda_\wedge})\lambda_\vee^{d/2} \prod_{i\in J} P\{X_i > (t_i - x_i^*)/\sqrt{\lambda_\wedge}\}.$$

Since further $x_i^* \geq t_i$, $i \in J$ the bound for $c_3(t)$ follows easily.

The other bound is obtained based on Jensen's inequality following closely the proof of Steck (1979). By (3.3) we have as in the proof of Proposition 3.1

$$\begin{aligned} P\{X > t\} &= \frac{\exp(-\alpha_t/2)}{(2\pi)^{d/2}|\Sigma|^{1/2}} \int_{x > t - x^*} \exp(-\langle x, Bx\rangle/2 - \langle x_I, \Sigma_{II}^{-1} t_I\rangle) dx \\ &= \frac{\exp(-\alpha_t/2)}{(2\pi)^{d/2}|\Sigma|^{1/2}} \\ &\quad \cdot \int_{x > t - x^*} \exp(-\langle x, \bar{B}x\rangle/2 - \langle x\sqrt{b}, x\sqrt{b}\rangle/2 - \langle x_I, \Sigma_{II}^{-1} t_I\rangle) dx, \end{aligned}$$

with $\boldsymbol{b} = (b_{11}, \ldots, b_{dd})'$, the main diagonal of the matrix B. Clearly by the positive definiteness of B we have $\boldsymbol{b} > \boldsymbol{0}$. Further $\bar{h}_i$ are positive for all $i \in I$ implying $R(\bar{h}_i) > 0$. Next, since

$$\begin{aligned}
&\int_{\boldsymbol{x}>\boldsymbol{t}-\boldsymbol{x}^*} \exp(-\langle \boldsymbol{x}\sqrt{\boldsymbol{b}}, \boldsymbol{x}\sqrt{\boldsymbol{b}}\rangle/2 - \langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle) d\boldsymbol{x}\\
&\quad = \int_{\boldsymbol{x}_J>(\boldsymbol{t}-\boldsymbol{x}^*)_J} \exp(-\langle(\boldsymbol{x}\sqrt{\boldsymbol{b}})_J, (\boldsymbol{x}\sqrt{\boldsymbol{b}})_J\rangle/2) d\boldsymbol{x}_J\\
&\quad \times \int_{(0,\infty)^{|I|}} \exp(-\langle(\boldsymbol{x}\sqrt{\boldsymbol{b}})_I, (\boldsymbol{x}\sqrt{\boldsymbol{b}})_I\rangle/2 - \langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle) d\boldsymbol{x}_I\\
&\quad = \frac{(2\pi)^{|J|/2}}{\prod_{i=1}^d \sqrt{b_{ii}}} \prod_{i\in J} \boldsymbol{P}\{X_i > (t - x_i^*)\sqrt{b_{ii}}\} \prod_{i\in I} R(\bar{h}_i)\\
&\quad =: p_t \geq \frac{(2\pi)^{|J|/2} 2^{-|J|}}{\prod_{i=1}^d \sqrt{b_{ii}}} \prod_{i\in I} R(\bar{h}_i) > 0
\end{aligned}$$

Jensen's inequality implies

$$\begin{aligned}
(2\pi)^{d/2}|\Sigma|^{1/2}\exp(\alpha_t/2)\boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}\} &= p_t \boldsymbol{E}\{\exp(-\langle \boldsymbol{U}, \bar{B}\boldsymbol{U}\rangle/2)\}\\
&\geq p_t \exp(-\langle \boldsymbol{E}\{\boldsymbol{U}\}, \bar{B}\boldsymbol{E}\{\boldsymbol{U}\}\rangle/2),
\end{aligned}$$

with $\boldsymbol{U} = (U_1, \ldots, U_d)$ a random vector in $\mathbb{R}^d$ with independent components and density function

$$p_t^{-1} \exp(-\langle \boldsymbol{x}\sqrt{\boldsymbol{b}}, \boldsymbol{x}\sqrt{\boldsymbol{b}}\rangle/2 - \langle \boldsymbol{x}_I, \Sigma_{II}^{-1}\boldsymbol{t}_I\rangle)\mathbf{1}_{\{\boldsymbol{x}_I>\boldsymbol{0}_I, \boldsymbol{x}_J>(\boldsymbol{t}-\boldsymbol{x}^*)_J\}}.$$

For $i \in I$ partial integration yields

$$\sqrt{b_{ii}}R(\bar{h}_i)\boldsymbol{E}\{U_i\} = \int_0^\infty s\exp(-s^2/2 - \bar{h}_i s)ds = 1 - \bar{h}_i R(\bar{h}_i),$$

and for $i \in J$ (recall $\boldsymbol{P}\{X_i > (t - x_i^*)\} \geq 1/2, i \in J$)

$$\begin{aligned}
\sqrt{b_{ii}}\boldsymbol{E}\{U_i\} &= \frac{\int_{(t-x_i^*)\sqrt{b_{ii}}}^\infty s\exp(-s^2/2)ds}{\sqrt{2\pi}\boldsymbol{P}\{X_i > (t - x_i^*)\sqrt{b_{ii}}\}} = \frac{\exp(-b_{ii}(t_i - x_i^*)^2/2)}{\sqrt{2\pi}\boldsymbol{P}\{X_i > (t - x_i^*)\sqrt{b_{ii}}\}}\\
&= \frac{1}{R((t - x_i^*)\sqrt{b_{ii}})},
\end{aligned}$$

hence the proof is complete. □

Remarks. a) In light of Proposition 2.1 if $|J^*| > 0$ then $t_i = x_i^*$ for all $i \in J^*$ hence

$$\boldsymbol{P}\{X_i > t_i - x_i^*\} = 1/2, \qquad V_i = 1/R(0) = \sqrt{2\pi}/2.$$

Further

$$\varphi(\boldsymbol{x}^*) = \frac{\exp(-\alpha_t/2)}{(2\pi)^{d/2}|\Sigma|^{1/2}}.$$

b) Clearly V_i tends to ∞ if $h_i \to \infty$ for $i \in I$, making the term $c_4(\boldsymbol{t})$ much smaller than $c_3(\boldsymbol{t})$.

c) For the asymptotic theory it is of great interest (cf. Bischoff *et al.* (2000, 2001)) to investigate the rate of convergence to 0 of $\boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}\}$ if one of the components $t_i \to \infty$. Both upper and lower bounds derived above capture very well the rate of convergence to 0 if all $h_i \to \infty$. We note that

$$\frac{\boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}\}}{\exp(-\alpha_t/2)\prod_{i\in I} R(h_i)} \tag{3.8}$$

is bounded away from 0 and ∞ when $\boldsymbol{t} \notin (-\infty, 0]^d$.

Next, we demonstrate the implications of the obtained upper and lower bounds by two examples.

Example 1. Let ξ, η be two Gaussian random variables with mean zero and unit standard deviation with correlation $\mathrm{Corr}\{\xi,\eta\} = \rho \in (-1,1)$. Further fix threshold $\boldsymbol{t} = t(b_1, b_2)$ with constants $b_1 > b_2 > 0$, $t > 0$. In view of (1.5) three cases have two be dealt with separately.

i) *Case* $\rho < b_2/b_1$. Simple calculations show that $I = \{1,2\}$, $J = J^* = \emptyset$ and

$$|\Sigma| = 1-\rho^2, \qquad h_1 = \frac{b_1 - \rho b_2}{1-\rho^2}, \qquad h_2 = \frac{b_2-\rho b_1}{1-\rho^2},$$
$$\lambda_1 = 1-\rho, \qquad \lambda_2 = 1+\rho, \qquad b_{11} = b_{22} = (1-\rho^2)^{-1},$$

$$\alpha_t = \langle (b_1,b_2), \Sigma^{-1}(b_1,b_2)\rangle = (b_1^2 + b_2^2 - 2\rho b_1 b_2)/(1-\rho^2) \geq (b_2-b_1)^2/(1-\rho^2) > 0$$

where Σ is the underlying correlation matrix of (ξ,η). Hence by the previous results we obtain

$$\boldsymbol{P}\{\xi > tb_1, \eta > tb_2\} \leq \frac{(1-\rho^2)^{3/2}\exp(-t^2(b_1^2+b_2^2-2\rho b_1 b_2)/(1-\rho^2)/2)}{2\pi(b_1-\rho b_2)(b_2-\rho b_1)}$$

and

$$\begin{aligned}\boldsymbol{P}\{\xi > tb_1, \eta > tb_2\} \geq \max(&r_1(b_1,b_2)r_1(b_2,b_1)\max(1-\rho, 1+\rho), (1-\rho^2)\\ &\times r_2(b_1,b_2)r_2(b_2,b_1)\\ &\times \exp(\rho/(2(1-\rho^2)))(r_3(b_1,b_2)+r_3(b_2,b_1)))),\end{aligned}$$

with

$$r_1(x,y) := R\left(\frac{(x-\rho y)\sqrt{\max(1-\rho,1+\rho)}}{1-\rho^2}\right), \qquad x,y \in I\!R$$

$$r_2(x,y) := R\left(\frac{x-\rho y}{\sqrt{1-\rho^2}}\right), \qquad r_3(x,y) := \sqrt{1-\rho^2}\left[r_2(x,y)^{-1} - \frac{x-\rho y}{1-\rho^2}\right], \qquad x,y \in I\!R.$$

ii) *Case* $\rho = b_2/b_1$. Solving the quadratic programming problem $\mathcal{P}_{B,(tb_1,tb_2)}$ we obtain

$$I = \{1\}, \qquad J = J^* = \{2\}, \qquad \Sigma_{II} = 1, \qquad \alpha_t = t^2 b_1^2, \qquad h_1 = tb_1.$$

Since $\rho > 0$ we get further

$$\lambda_\wedge = 1-\rho, \qquad \lambda_\vee = 1+\rho.$$

So we have

$$P\{\xi > tb_1, \eta > tb_2\} \le \frac{\exp(-t^2 b_1^2/2)}{\sqrt{2\pi(1-\rho^2)}} \min\left(\frac{\sqrt{1-\rho^2}}{tb_1}, \frac{(1+\rho)R(tb_1\sqrt{1+\rho})}{2}, \frac{1}{2tb_1}\right)$$

$$P\{\xi > tb_1, \eta > tb_2\} \ge \frac{\exp(-t^2 b_1^2/2)}{\sqrt{2\pi(1-\rho^2)}} \max\Bigg(\frac{(1-\rho)R(tb_1\sqrt{(1-\rho)})}{2}, \frac{(1-\rho^2)R(tb_1\sqrt{1-\rho^2})}{2} \times \exp(\rho/(2\sqrt{1-\rho^2}) \times [\sqrt{2\pi}/2 + 1/R(tb_1\sqrt{1-\rho^2}) - tb_1\sqrt{1-\rho^2}])\Bigg).$$

iii) *Case* $\rho \in (b_2/b_1, 1)$. The only difference with the previous case is that $J^* = \emptyset$. It can be easily seen that $\boldsymbol{x}^* = t(b_1, \rho b_1)$, and further α_t, h_1, I, J are the same as in the previous case. To obtain an explicit expression for the upper and lower bounds instead of $1/2$ corresponding to $P\{X_1 > 0\}$ we should put now in the above inequalities $P\{X_1 > t(b_2 - \rho b_1)\} > 1/2$ if $t > 0$. Observe that in the above example we have $0 < c_0(\boldsymbol{t}) < c_2(\boldsymbol{t})$ for t large.

Example 2. Proposition 4.2 of Raab (1999) estimate the tail probability of Gaussian random vectors needed for a compound Poisson approximation problem related to m-dependent stationary Gaussian random sequences. Our example gives a stronger result: let $\boldsymbol{t}_n$, $n \ge 1$ be such that

$$\boldsymbol{t}_n = \mathbf{1} b_n + o(\mathbf{1})/b_n,$$

with b_n defined as the solution of $b_n = n\varphi(b_n)$ where φ denotes the standard normal density function and $o(\mathbf{1})$ a sequence of vectors in $\mathbb{R}^d$ vanishing to 0 as $n \to \infty$. It is well known (see e.g. Reiss (1989) or Falk *et al.* (1994)) that we may define for large n

$$b_n := \sqrt{2\ln n} - \frac{\ln(4\pi \ln n)}{2\sqrt{2\ln n}}.$$

Further it follows $I_n = I$ for n large and $\alpha_{t_n} = b_n^2 \langle \mathbf{1}_I, \Sigma_{II}^{-1}\mathbf{1}_I\rangle + o(1)$ as $n \to \infty$, hence we obtain by the previous results

$$0 < K_1 \le \frac{P\{\boldsymbol{X} > \mathbf{1}b_n + o(\mathbf{1})/b_n\}}{b_n^{-|I|}\exp(-b_n^2\langle \mathbf{1}_I, \Sigma_{II}^{-1}\mathbf{1}_I\rangle/2)} \le K_2 < \infty, \qquad \text{for} \quad n \to \infty$$

with K_1, K_2 two constants depending on Σ but not on n.

4. Asymptotic results

Let $\{\boldsymbol{X}_n, n \ge 1\}$ be a sequence of d-dimensional standard Gaussian random vectors with positive definite correlation matrix Σ_n. In this short section we treat the asymptotic behaviour of $P\{\boldsymbol{X}_n \ge \boldsymbol{t}_n\}$, with $\boldsymbol{t}_n$ so that at least one of its components tends to ∞. Again we rely on Proposition 2.1 considering now the quadratic programming problem $(\mathcal{P}_{B_n,\boldsymbol{t}_n})$, with B_n the inverse matrix of Σ_n and let $\alpha_{\boldsymbol{t}_n}$, I_n, J_n, J_n^*, $\boldsymbol{x}_n^*$ as in the aforementioned proposition. In the following we use a simplified notation; for example we write $\boldsymbol{t}_{n,I_n}$ instead of $(\boldsymbol{t}_n)_{I_n}$ and similarly for other cases. It does not lead to ambiguous interpretations. We may state now the main result of this section.

THEOREM 4.1. *Under the above setup suppose that for all large n we have $I_n = I$ and*

$$\lim_{n\to\infty}(\boldsymbol{t}_n - \boldsymbol{x}_n^*)_J = \boldsymbol{t}_J^*, \quad \text{and} \quad \forall i \in I : \lim_{n\to\infty}\langle \boldsymbol{e}_i, \Sigma_{n,II}^{-1}\boldsymbol{t}_{n,I}\rangle = \infty. \tag{4.1}$$

Suppose further that the sequence of matrices $B_n, n \in \mathbb{N}$, is bounded for all n and

$$\lim_{n\to\infty} B_{n,JJ} = B_{JJ}, \qquad \lim_{n\to\infty}\Sigma_{n,JJ}^{-1} = \Sigma_{JJ}^{-1}, \tag{4.2}$$

with B_{JJ}, Σ_{JJ} positive definite matrices. Then as $n \to \infty$

$$\boldsymbol{P}\{\boldsymbol{X}_n > \boldsymbol{t}_n\} = \frac{(1+o(1))\exp(-\alpha_{\boldsymbol{t}_n}/2)\boldsymbol{P}\{\boldsymbol{X}_{n,J} \ge \boldsymbol{t}_J^* \mid \boldsymbol{X}_{n,I} = \boldsymbol{0}_I\}}{(2\pi)^{|I|/2}|\Sigma_{n,II}|^{1/2}\prod_{i\in I}\langle \boldsymbol{e}_i, \Sigma_{n,II}^{-1}\boldsymbol{t}_{n,I}\rangle} \tag{4.3}$$

with $\boldsymbol{t}_J^* \le \boldsymbol{0}_J$ and

$$\alpha_{\boldsymbol{t}_n} = \min_{\boldsymbol{x}\ge\boldsymbol{t}_n}\langle \boldsymbol{x}, B_n\boldsymbol{x}\rangle \to \infty, \qquad n \to \infty.$$

PROOF. Assume in the following w.l.o.g. that $I_n = I$ holds for all $n \in \mathbb{N}$. Define $\tilde{\boldsymbol{x}}_n$ for given $\boldsymbol{x}$ by $(\tilde{\boldsymbol{x}}_n)_i := \langle \boldsymbol{e}_i, \Sigma_{n,II}^{-1}\boldsymbol{t}_{n,I}\rangle^{-1}x_i$, $i \in I$ and $(\tilde{\boldsymbol{x}}_n)_J := \boldsymbol{x}_J$. Using (3.4) we get

$$\varphi(\tilde{\boldsymbol{x}}_n + \boldsymbol{x}^*) = \frac{\exp(-\alpha_{\boldsymbol{t}_n}/2)}{(2\pi)^{d/2}|\Sigma_n|^{1/2}}\exp(-\langle \tilde{\boldsymbol{x}}_n, B_n\tilde{\boldsymbol{x}}_n\rangle/2 - \langle \boldsymbol{x}, \boldsymbol{1}_I\rangle),$$

and changing the variables yields (recall $\langle \boldsymbol{e}_i, \Sigma_{n,II}^{-1}\boldsymbol{t}_{n,I}\rangle > 0$ for all $i \in I$)

$$\boldsymbol{P}\{\boldsymbol{X}_n > \boldsymbol{t}_n\} = \left(\prod_{i\in I}\langle \boldsymbol{e}_i, \Sigma_{n,II}^{-1}\boldsymbol{t}_{n,I}\rangle^{-1}\right)\frac{\exp(-\alpha_{\boldsymbol{t}_n}/2)}{(2\pi)^{d/2}|\Sigma_n|^{1/2}}\mathcal{J}_n,$$

with

$$\mathcal{J}_n := \int_{\boldsymbol{x}>\boldsymbol{t}_n-\boldsymbol{x}_n^*}\exp(-\langle \tilde{\boldsymbol{x}}_n, B_n\tilde{\boldsymbol{x}}_n\rangle/2 - \langle \boldsymbol{x}, \boldsymbol{1}_I\rangle)d\boldsymbol{x}.$$

Since B_n is bounded and by (4.1) and (4.2), for every $\boldsymbol{x}$

$$\lim_{n\to\infty}\langle \tilde{\boldsymbol{x}}_n, B_n\tilde{\boldsymbol{x}}_n\rangle = \langle \boldsymbol{x}_J, B_{JJ}\boldsymbol{x}_J\rangle > 0$$

and similarly to (3.6) by condition (4.2)

$$\exp(-\langle \tilde{\boldsymbol{x}}_n, B_n\tilde{\boldsymbol{x}}_n\rangle/2 - \langle \boldsymbol{x}, \boldsymbol{1}_I\rangle) \le \exp(-\langle \boldsymbol{x}_J, \Sigma_{n,JJ}^{-1}\boldsymbol{x}_J\rangle/2) \to \exp(-\langle \boldsymbol{x}_J, \Sigma_{JJ}^{-1}\boldsymbol{x}_J\rangle/2)$$

for $n \to \infty$. Applying Lebesgue dominated convergence theorem (cf. Theorem 1.21 in Kallenberg (1997)) implies

$$\lim_{n\to\infty}\mathcal{J}_n = \int_{\boldsymbol{x}_J\ge\boldsymbol{t}_J^*}\exp(-\langle \boldsymbol{x}_J, B_{JJ}\boldsymbol{x}_J\rangle/2)d\boldsymbol{x}_J,$$

hence as $n \to \infty$ using (4.2)

$$\begin{aligned}\mathcal{J}_n &= (1+o(1))(2\pi)^{|J|/2}|B_{JJ}|^{-1/2}\int_{\boldsymbol{x}_J\geq \boldsymbol{t}_J^*}\frac{\exp(-\langle \boldsymbol{x}_J, B_{JJ}\boldsymbol{x}_J\rangle/2)}{(2\pi)^{|J|/2}|B_{JJ}|^{-1/2}}d\boldsymbol{x}_J\\ &= (1+o(1))(2\pi)^{|J|/2}|B_{n,JJ}|^{-1/2}\boldsymbol{P}\{\boldsymbol{X}_{n,J} > \boldsymbol{t}_J^* \mid \boldsymbol{X}_{n,I} = \boldsymbol{0}_I\}\\ &= (1+o(1))(2\pi)^{|J|/2}|\Sigma_{n,II}|^{-1/2}|\Sigma_n|^{1/2}\boldsymbol{P}\{\boldsymbol{X}_{n,J} > \boldsymbol{t}_J^* \mid \boldsymbol{X}_{n,I} = \boldsymbol{0}_I\},\end{aligned}$$

and the fact that

$$|\Sigma_n| = |\Sigma_{n,II}||\Sigma_{n,JJ} - \Sigma_{n,JI}\Sigma_{n,II}^{-1}\Sigma_{n,IJ}| = |\Sigma_{n,II}||B_{n,JJ}^{-1}| > 0$$

completes the proof. □

Remarks. a) In the above theorem for J^* not empty we have $\boldsymbol{t}_{J^*}^* = \boldsymbol{0}_{J^*}$ and $\boldsymbol{t}_{J\setminus J^*}^* \in [-\infty, 0]^{|J\setminus J^*|}$.

b) For the term in the asymptotic result above we may write further for $n \geq 2$

$$\frac{\exp(-\alpha_{t_n}/2)}{(2\pi)^{|I|/2}|\Sigma_{n,II}|^{1/2}} = \frac{\exp(-\langle \boldsymbol{t}_{n,I}, \Sigma_{n,II}^{-1}\boldsymbol{t}_{n,I}\rangle/2)}{(2\pi)^{|I|/2}|\Sigma_{n,II}|^{1/2}} = \varphi_I(\boldsymbol{x}_I^*) = \varphi_I(\boldsymbol{t}_I),$$

with φ_I the density function of $\boldsymbol{X}_{n,I}$.

Simple thresholds are treated in the next corollary.

COROLLARY 4.1. *Let $\boldsymbol{X}$ be a standard Gaussian random vector on $\mathbb{R}^d$, $d \geq 2$, with correlation matrix Σ and $\boldsymbol{t}_n = t_n\boldsymbol{c}$, $t_n > 0$, $\boldsymbol{c} \notin (-\infty, 0]^d$. Then as $n \to \infty$*

$$\boldsymbol{P}\{\boldsymbol{X} > \boldsymbol{t}_n\} = \frac{(1+o(1))\exp(-t_n^2\langle \boldsymbol{c}_I, \Sigma_{II}^{-1}\boldsymbol{c}_I\rangle/2)\boldsymbol{P}\{\boldsymbol{X}_{J^*} \geq \boldsymbol{0}_{J^*} \mid \boldsymbol{X}_I = \boldsymbol{0}_I\}}{(2\pi)^{|I|/2}|\Sigma_{II}|^{1/2}t_n^{|I|}\prod_{i\in I}h_i} \tag{4.4}$$

holds as $n \to \infty$, with $I, J^*, h_i := \langle \boldsymbol{e}_i, \Sigma_{II}^{-1}\boldsymbol{c}_I\rangle > 0$, $i \in I$ defined with respect to solution of quadratic programming problem $\mathcal{P}_{B,\boldsymbol{c}}$.

PROOF. Let I_n, J_n^* and I, J^* be the index set related to the quadratic programming problems $\mathcal{P}_{B,\boldsymbol{t}_n}$ and $\mathcal{P}_{B,\boldsymbol{c}}$ respectively. It follows that $I_n = I$, $J_n^* = J^*$ and the unique solution of $\mathcal{P}_{B,\boldsymbol{t}_n}$ is $\boldsymbol{x}^* = t_n\boldsymbol{c}^*$ with $\boldsymbol{c}^*$ the unique solution of $\mathcal{P}_{B,\boldsymbol{c}}$. Further for all $i \in I$ we have as $n \to \infty$

$$\langle \boldsymbol{e}_i, \Sigma_{II}^{-1}\boldsymbol{t}_{n,I}\rangle = t_n\langle \boldsymbol{e}_i, \Sigma_{II}^{-1}\boldsymbol{c}_I\rangle \to \infty.$$

Next, by Proposition 2.1

$$\boldsymbol{c}_{J^*}^* = \boldsymbol{c}_{J^*}, \qquad \boldsymbol{c}_{J\setminus J^*}^* > \boldsymbol{c}_{J\setminus J^*}$$

(whenever the index sets above are not empty), hence as $n \to \infty$

$$\boldsymbol{x}_{n,J\setminus J^*}^* = t_n(\boldsymbol{c} - \boldsymbol{c}^*)_{J\setminus J^*} \to (-\infty, \ldots, -\infty), \qquad \text{and} \qquad \boldsymbol{x}_{n,J^*}^* = \boldsymbol{0}_{J^*}.$$

Since for all $n \in \mathbb{N}$

$$\alpha_{t_n} = t_n^2\langle \boldsymbol{c}_I^*, \Sigma_{II}^{-1}\boldsymbol{c}^*\rangle = t_n^2\langle \boldsymbol{c}_I, \Sigma_{II}^{-1}\boldsymbol{c}_I\rangle > 0$$

the proof follows immediately from the previous theorem. □

Finally, we discuss a particular situation, where the asymptotic expression does not depend on the solution of the quadratic programming problem $(\mathcal{P}_{B_n,t_n})$.

Consider $\boldsymbol{X}_n$ as above and threshold

$$t_n = a_n \boldsymbol{x} + b_n \mathbf{1}, \qquad \boldsymbol{x} \in I\!R^d, \tag{4.5}$$

with a_n, b_n positive constants. Assuming that

$$a_n^2 B_n = a_n^2 \Sigma_n^{-1} \to \Sigma_a^{-1}, \qquad n \to \infty,$$

with Σ_a some positive definite matrix, and for all $i = 1, \ldots, d$

$$\lim_{n\to\infty} a_n b_n \langle e_i, B_n \mathbf{1} \rangle = h_i$$

we get

$$\begin{aligned} P\{\boldsymbol{X}_n > \boldsymbol{t}_n\} &= \frac{a_n^d \exp(-\langle \boldsymbol{t}_n, B_n \boldsymbol{t}_n \rangle/2)}{(2\pi)^{d/2} |\Sigma_n|^{1/2}} \int_{\boldsymbol{y}>\boldsymbol{x}} \exp(-a_n^2 \langle \boldsymbol{y}, B_n \boldsymbol{y} \rangle/2 - a_n b_n \langle \boldsymbol{y}, B_n \mathbf{1} \rangle) d\boldsymbol{y} \\ &= \frac{(1+o(1)) a_n^d \exp(-\langle \boldsymbol{t}_n, B_n \boldsymbol{t}_n \rangle/2)}{(2\pi)^{d/2} |\Sigma_n|^{1/2}} \\ &\quad \times \int_{\boldsymbol{y}>\boldsymbol{x}} \exp\left(-\langle \boldsymbol{y}, \Sigma_a^{-1} \boldsymbol{y} \rangle/2 - \sum_{i=1}^d h_i y_i\right) d\boldsymbol{y}, \end{aligned}$$

which leads to another proof of Theorem 2 of Hüsler and Reiss (1989).

5. Comments and examples

The asymptotic results obtained in the previous section open the way for treatment of more difficult situations when Gaussian random processes are involved. These ideas are successfully developed in Bischoff *et al.* (2000, 2001).

Our results are of certain important for the extreme value theory and rare events for Gaussian triangular arrays. In extreme value theory usually threshold of the form $\boldsymbol{a}_n \boldsymbol{x} + \boldsymbol{b}_n$ are considered.

Our last comment is in connection with the speed of convergence in (4.3). Clearly when α_{t_n} goes to infinity, the speed of convergence to 0 is exponential fast. Now if $h_i = \langle e_i, \Sigma_{n,II}^{-1} \boldsymbol{t}_{n,I} \rangle \to \infty$ then at least for one i we have $t_{ni} \to \infty$. It is easy to see (splitting the scalar product as in (3.5)) that

$$\alpha_{t_n} = \min_{\boldsymbol{x} \geq \boldsymbol{t}_n} \langle \boldsymbol{x}, B_n \boldsymbol{x} \rangle \geq \lambda_n^{-1} \max_{i=1,\ldots,d} t_{ni}^2,$$

with λ_n^{-1} the maximal eigenvalue of B_n^{-1}. So under conditions of the Theorem 4.1 the speed of convergence is exponentially fast since

$$\lim_{n\to\infty} \alpha_{t_n} = \infty$$

follows easily by the above arguments.

It is clear that there is an intrinsic relation between the asymptotic behaviour of the components of $\boldsymbol{t}_n$ and the index set I_n. If one of the components say for instance the first dominates all the others, i.e.

$$\lim_{n\to\infty} t_{ni}/t_{n1} = 0,$$

with $t_{n1} \to \infty$, it follows that $I_n = \{1\}$ for all n large, so $\alpha_{t_n} = t_{n1}^2 + o(1)$ and $h_1 = t_{n1}$, hence the contribution of the other components is negligible for the asymptotic expansion.

Finally we consider three examples illustrating the results for the asymptotic analysis.

Example 3. Let $(\xi_1, \eta_1), \ldots, (\xi_n, \eta_n)$ iid random vectors with standard Gaussian random components. Assume that

$$\mathrm{Corr}\{\xi_k, \eta_k\} = \rho_k \to \rho \in (-1, 1), \qquad k \to \infty$$

and take $\boldsymbol{t}_n = t_n(b_1, b_2)$ as in Example 1. By the above results if $\rho < \min(b_1/b_2, b_2/b_1)$, we have as $n \to \infty$

$$\boldsymbol{P}\{\xi_n > t_n b_1, \eta_n > t_n b_2\} = (1 + o(1)) \frac{(1-\rho^2)^{3/2} \exp(-t_n^2 \alpha/2)}{2\pi(b_1 - \rho b_2)(b_2 - \rho b_1) t_n^2}$$

with

$$\alpha = [b_1^2 - 2\rho b_1 b_2 + b_2^2]/(1 - \rho^2).$$

If $\rho_n = b_2/b_1$ for all n large, we get $J_n = J_n^* = \{2\}$, hence in view of Theorem 4.1

$$\boldsymbol{P}\{\xi_n > t_n b_1, \eta_n > t_n b_2\} = (1 + o(1)) \frac{\exp(-t_n^2 b_1^2/2)}{2 t_n b_1 \sqrt{2\pi}}, \qquad n \to \infty.$$

It is easy to see that if $\limsup_{n\to\infty} \rho_n > b_2/b_1 > 0$ we get $J_n = \{2\}$, $|J_n^*| = 0$, thus we obtain by the same theorem

$$\boldsymbol{P}\{\xi_n > t_n b_1, \eta_n > t_n b_2\} = (1 + o(1)) \frac{\exp(-t_n^2 b_1^2)}{t_n b_1 \sqrt{2\pi}}, \qquad n \to \infty.$$

Note that in both last cases the asymptotic does not depend on ρ_n.

Alternatively, since

$$\begin{aligned}\boldsymbol{P}&\{\xi_n > t_n b_1, \eta_n > t_n b_2\}\\ &= \frac{1}{\sqrt{2\pi}} \int_0^\infty \varphi(x + b_1 t_n)(1 - \Phi(([b_2 - \rho_n b_1] t_n - x\rho_n)/\sqrt{1 - \rho_n^2})) dx\end{aligned}$$

Watson Lemma establishes the asymptotics if $\rho_n = \rho$, $n \in \mathbb{N}$, which is also given in Lemma 2.1 of Elnaggar and Mukherjea (1999). For the case $\rho = b_2/b_1$ the asymptotic expression is obtained therein up to some constant.

Example 4. Let us consider the case $\Sigma_n \in \mathbb{R}^{d \times d}$, $d \geq 2$, is permutation symmetric, i.e.

$$\Sigma_n = (1 - \rho_n)\mathcal{I}_d + \rho_n \mathbf{1}\mathbf{1}', \tag{5.1}$$

with $-1/(d-1) < \rho_n < 1$, $n \in \mathbb{N}$ and $\mathcal{I}_d$ the identity matrix. The inverse matrix is then

$$B_n = \frac{1}{1 - \rho_n} \mathcal{I}_d - \frac{\rho_n}{(1 - \rho_n)(1 + (d-1)\rho_n)} \mathbf{1}\mathbf{1}',$$

hence as $n \to \infty$

$$\langle e_i, B_n \mathbf{1} \rangle = \frac{1}{1+(d-1)\rho_n} \to \frac{1}{1+(d-1)\rho} > 0,$$

iff

$$\lim_{n\to\infty} \rho_n = \rho \in (-1/(d-1), 1].$$

In order to make sure that the inverse matrix B_n remains bounded for all n large (required in Theorem 4.1) we need to exclude $\rho = 1$.

If $\boldsymbol{t_n} = (1+o(1))t_n\mathbf{1}$ then $|I_n| = d$, hence we have

$$\boldsymbol{P}\{\boldsymbol{X_n} > \boldsymbol{t_n}\} = \varphi(t_n\mathbf{1})[(1+(d-1)\rho)/t_n]^d, \qquad \text{as} \quad n \to \infty, \tag{5.2}$$

which reduces for $\rho = 0$ to Lemma D in Shao and Santosh (1999), whereas for $t_n = b_n + o(1/b_n)$ and b_n as in Example 2 the above results is stronger than the claim of Proposition 4.2 of Raab (1999) dealing only with the case $d = 3$.

Example 5. Consider now the correlation matrix similar to the example of Gnedin (1998)

$$\Sigma_n = \begin{pmatrix} 1 & \rho_n & \rho_n \\ \rho_n & 1 & 0 \\ \rho_n & 0 & 1 \end{pmatrix}, \qquad B_n = \frac{1}{1-2\rho_n^2}\begin{pmatrix} 1 & -\rho_n & -\rho_n \\ -\rho_n & 1-\rho_n^2 & \rho_n^2 \\ -\rho_n & \rho_n^2 & 1-\rho_n^2 \end{pmatrix}$$

so that

$$\lim_{n\to\infty} \rho_n = \rho \in (-1/\sqrt{2}, 1/\sqrt{2}).$$

Take threshold $\boldsymbol{t_n} = t_n(1,1,1)'$ with $\lim_{n\to\infty} t_n = \infty$. The interesting cases when our results can be applied are $\rho_n = 1/2$ for large n or $\limsup_{n\to\infty} \rho_n \in (1/2, 1/\sqrt{2})$, since the first row of the matrix B_n sums to some nonpositive constant, implying that (1.4) does not hold.

It is easily seen that $J_n = \{1\}$ and Σ_{II} is the identity matrix in $\mathbb{R}^2$. In the first case $J_n^* = J_n$ hence

$$\boldsymbol{P}\{\boldsymbol{X_n} > t_n(1,1,1)'\} = (1+o(1))\exp(-t_n^2/2)/(4\pi t_n^2), \qquad \text{as} \quad n \to \infty$$

and for the second case

$$\boldsymbol{P}\{\boldsymbol{X_n} > t_n(1,1,1)'\} = (1+o(1))\exp(-t_n^2/2)/(2\pi t_n^2), \qquad \text{as} \quad n \to \infty,$$

since we have further $J_n^* = \emptyset$.

Acknowledgements

We thank Vladimir I. Piterbarg and Alexander V. Gnedin for useful hints and an anonymous referee for very helpful suggestions.

References

Barndorff-Nielsen, O. E. and Cox, D. R. (1989). *Asymptotic Techniques for Use in Statistics*, Chapman & Hall, London.

Bischoff, W., Hashorva, E., Hüsler, J. and Miller. F. (2000). Asymptotics of the extremes of Gaussian processes with trend with applications to the power of the Kolmogoroff test for several problems (preprint).

Bischoff, W., Hashorva, E., Hüsler, J. and Miller, F. (2001). Asymptotically optimal test for a change-point regression problem with application in quality control (preprint).

Elnaggar, M. and Mukherjea, A. (1999). Identification of the parameters of a trivariate normal vector by the distribution of the minimum, *J. Statist. Plann. Inference*, **78**, 23–37.

Falk, M., Hüsler, J. and Reiss, R. D. (1994). *Laws of Small Numbers: Extremes and Rare Events*, DMV Seminar, Vol. 23, Birkhäuser, Basel.

Fang, K.-T., Xu, J.-L. (1990). The Mills' ratio of multivariate normal distributions and spherical distributions, *Statistical Inference in Elliptically Contoured and Related Distributions*, 457–468, Allerton, New York.

Gjacjauskas, E. (1973). Estimation of the multidimensional normal probability distribution law for a receding hyperangle, *Litovsk. Mat. Sb.*, **13**(3), 83–90.

Gnedin, A. V. (1998). Records from a multivariate normal sample, *Statist. Probab. Lett.*, **39**, 11–15.

Hashorva, E. and Hüsler, J. (2002). Exact asymptotics for records of Gaussian random vectors, *Stochastic Models*, **18**(1), 41–69.

Hüsler, J. and Reiss, R. D. (1989). Maxima of normal random vectors: Between independence and complete dependence, *Statist. Probab. Lett.*, **7**, 283–286.

Kallenberg, O. (1997). *Foundations of Modern Probability*, Springer, New York.

Raab, M. (1999). Compound Poisson approximation of the number of exceedances in Gaussian sequences, *Extremes*, **1**(3), 295–321.

Reiss, R.-D. (1989). *Approximate Distributions of Order Statistics: With Applications to Nonparametric Statistics*, Springer, New York.

Satish, I. (1986). On a lower bound for the multivariate normal Mills' ratio, *Ann. Probab.*, **14**, 1399–1403.

Savage, I. R. (1962). Mills' ratio for multivariate normal distribution, *J. Res. Nat. Bur. Standards Sect. B*, **66**, 93–96.

Shao, C. F. and Santosh, S. V. (1999). Learning finite binary sequences from half-space data, *Random Structures Algorithms*, **14**(4), 345–381.

Stanislaw, J. S. and Werner, E. (1999). A nonsymmetric correlation inequality for Gaussian measure, *J. Multivariate Anal.*, **68**, 193–211.

Steck, G. P. (1979). Lower bounds for the multivariate normal Mills' ratio, *Ann. Probab.*, **7**, 547–551.

Tong, Y. L. (1989). *The Multivariate Normal Distribution*, Springer, Berlin.