ACTA ET COMMENTATIONES UNIVERSITATIS TARTUENSIS DE MATHEMATICA
Volume 11, 2007

On the joint distribution of a linear and a quadratic form in skew normal variables

Arjun K. Gupta, Tõnu Kollo, and Anne Selart

Abstract. Let $\mathbf{z}$ be distributed as multivariate skew normal vector. We derive the joint moment generating function (m.g.f.) of a linear form and a quadratic form in $\mathbf{z}$, and the conditions for their independence. First two multivariate cumulants of the two forms are derived and applied in special cases. Finally a simulation example is presented.

1. Introduction

The multivariate skew normal distribution has been introduced in Azzalini and Dalla Valle (1996) with the first applications given in Azzalini and Capitanio (1999). This class of distributions includes the normal family and has some properties like the normal, and yet is skew. It is useful in robustness studies and appears in the theory of linear models related to hidden truncation and selective reporting (Arnold and Beaver, 2002; Arellano-Valle et al., 2006). Following Gupta and Kollo (2003) the random p-vector $\mathbf{z}$ has a multivariate skew normal distribution if it is continuous and its probability density function (p.d.f.) is given by

$$f_{\mathbf{z}}(\mathbf{z}) = 2\phi_p(\mathbf{z};\Sigma)\Phi(\boldsymbol{\alpha}'\mathbf{z}), \quad \mathbf{z} \in \mathbb{R}^p, \tag{1.1}$$

where $\Sigma : p \times p$ is positive definite (p.d.), $\boldsymbol{\alpha} \in \mathbb{R}^p$, $\phi_p(\mathbf{z};\Sigma)$ is the density of $N_p(\mathbf{0},\Sigma)$ and $\Phi(\cdot)$ is the cumulative distribution function (c.d.f.) of $N(0,1)$. We will denote by $\mathbf{z} \sim SN_p(\Sigma,\boldsymbol{\alpha})$, to mean that the random vector $\mathbf{z}$ has p-variate skew normal density (1.1). The m.g.f. of $\mathbf{z} \sim SN_p(\Sigma,\boldsymbol{\alpha})$ is

$$M(\mathbf{t}) = 2e^{\frac{1}{2}\mathbf{t}'\Sigma\mathbf{t}}\Phi\left(\frac{\boldsymbol{\alpha}'\Sigma\mathbf{t}}{(1+\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha})^{\frac{1}{2}}}\right), \quad \mathbf{t} \in \mathbb{R}^p. \tag{1.2}$$

Received December 10, 2007.
2000 *Mathematics Subject Classification.* Primary 62H10; Secondary 62E15.
Key words and phrases. Cumulants, linear form, moments, moment generating function, quadratic form, skew normal distribution.

The mean vector and the covariance matrix of $\mathbf{z}$ are given by

$$\mu_{\mathbf{z}} = E(\mathbf{z}) = \sqrt{\frac{2}{\pi}}\boldsymbol{\delta},$$

$$\operatorname{cov}(\mathbf{z}) = \Sigma - \mu_{\mathbf{z}}\mu_{\mathbf{z}}',$$

where $\boldsymbol{\delta} = (1 + \boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha})^{-\frac{1}{2}}\Sigma\boldsymbol{\alpha}$. Note that the mean value given in Azzalini and Capitanio (1999) is in error.

Many problems in statistics require the knowledge of the joint distribution of a linear form and a quadratic form of a given random vector, e.g. in simultaneous confidence interval construction or a simultaneous control chart studies (e.g. see Schöne, 1997; Schöne and Schmid, 2000; Knoth et al., 2001).

In this paper, we derive the joint m.g.f. of a linear and a quadratic form for a skew normal vector $\mathbf{z}$ in Section 2. In Section 3 conditions for independence of the two forms are derived. Section 4 deals with finding the first two cumulants of a linear form and a quadratic form in $\mathbf{z}$. The paper is concluded by an example based on simulations.

2. Joint m.g.f. of a linear and a quadratic form

In this section we derive the joint m.g.f. of a linear form and a quadratic form of a skew normal random vector. Let $\mathbf{z} \sim SN_p(\Sigma, \boldsymbol{\alpha})$. Consider the linear form $B\mathbf{z}$ where B is $q \times p$ and the quadratic form $\mathbf{z}'A\mathbf{z}$ where $A' = A$. First we state a lemma (see Zacks, 1981, pp. 53-54) which is needed in the sequel.

Lemma 1. *Let $\mathbf{y} \sim N_p(\mathbf{0}, \Sigma)$. Then, for any scalar u and vector $\mathbf{v} \in \mathbb{R}^p$, we have*

$$E[\Phi(u + \mathbf{v}'\mathbf{y})] = \Phi\left(\frac{u}{(1 + \mathbf{v}'\Sigma\mathbf{v})^{\frac{1}{2}}}\right).$$

In the next theorem we give the joint m.g.f. of $B\mathbf{z}$ and $\mathbf{z}'A\mathbf{z}$.

Theorem 1. *Let B be a $q \times p$-matrix and $A : p \times p$ symmetric matrix. If $\mathbf{z} \sim SN_p(\Sigma, \boldsymbol{\alpha})$ then the joint m.g.f. of $B\mathbf{z}$ and $\mathbf{z}'A\mathbf{z}$ is*

$$M(t_0, \mathbf{t}) = \frac{2e^{\frac{1}{2}\mathbf{t}'B(\Sigma^{-1}-2t_0A)^{-1}B'\mathbf{t}}}{|I_p - 2t_0A\Sigma|^{\frac{1}{2}}}\,\Phi\left(\frac{\boldsymbol{\alpha}'(\Sigma^{-1} - 2t_0A)^{-1}B'\mathbf{t}}{(1 + \boldsymbol{\alpha}'(\Sigma^{-1} - 2t_0A)^{-1}\boldsymbol{\alpha})^{\frac{1}{2}}}\right). \quad (2.1)$$

Proof. For $t_0 \in \mathbb{R}$ and $\mathbf{t} \in \mathbb{R}^q$ we have

$$M(t_0, \mathbf{t}) = 2\int_{\mathbb{R}^p} \exp(t_0\mathbf{z}'A\mathbf{z} + \mathbf{t}'B\mathbf{z})\phi_p(\mathbf{z}; \Sigma)\Phi(\boldsymbol{\alpha}'\mathbf{z})d\mathbf{z}$$

$$= 2\int_{\mathbb{R}^p} \exp\left(-\frac{1}{2}(\mathbf{z}'(\Sigma^{-1} - 2t_0A)\mathbf{z} - 2\mathbf{t}'B\mathbf{z})\right)\frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}}\Phi(\boldsymbol{\alpha}'\mathbf{z})d\mathbf{z}.$$

Let us apply change of variable

$$\mathbf{y} = (\Sigma^{-1} - 2t_0 A)^{\frac{1}{2}}\mathbf{z} - (\Sigma^{-1} - 2t_0 A)^{-\frac{1}{2}}B'\mathbf{t}.$$

Then

$$\mathbf{z} = (\Sigma^{-1} - 2t_0 A)^{-\frac{1}{2}}\mathbf{y} + (\Sigma^{-1} - 2t_0 A)^{-1}B'\mathbf{t}$$

and

$$\mathbf{y}'\mathbf{y} = \mathbf{z}'(\Sigma^{-1} - 2t_0 A)\mathbf{z} + \mathbf{t}'B(\Sigma^{-1} - 2t_0 A)^{-1}B'\mathbf{t} - 2\mathbf{t}'B\mathbf{z}.$$

The Jacobian equals

$$\text{abs}\left(\left|\frac{d\mathbf{z}}{d\mathbf{y}}\right|\right) = |\Sigma^{-1} - 2t_0 A|^{-\frac{1}{2}}$$

and we have the representation

$$\begin{aligned} M(t_0, \mathbf{t}) &= \frac{2e^{\frac{1}{2}\mathbf{t}'B(\Sigma^{-1}-2t_0A)^{-1}B'\mathbf{t}}}{(2\pi)^{\frac{p}{2}}|I_p - 2t_0 A\Sigma|^{\frac{1}{2}}} \\ &\quad \times \int_{\mathbb{R}^p} e^{\frac{1}{2}\mathbf{y}'\mathbf{y}}\Phi(\boldsymbol{\alpha}'(\Sigma^{-1} - 2t_0 A)^{-1}B'\mathbf{t} + \boldsymbol{\alpha}'(\Sigma^{-1} - 2t_0 A)^{-\frac{1}{2}}\mathbf{y})d\mathbf{y} \\ &= \frac{2e^{\frac{1}{2}\mathbf{t}'B(\Sigma^{-1}-2t_0A)^{-1}B'\mathbf{t}}}{|I_p - 2t_0 A\Sigma|^{\frac{1}{2}}} \\ &\quad \times E_{\mathbf{y}}\Phi(\boldsymbol{\alpha}'(\Sigma^{-1} - 2t_0 A)^{-1}B'\mathbf{t} + \boldsymbol{\alpha}'(\Sigma^{-1} - 2t_0 A)^{-\frac{1}{2}}\mathbf{y}) \\ &= \frac{2e^{\frac{1}{2}\mathbf{t}'B(\Sigma^{-1}-2t_0A)^{-1}B'\mathbf{t}}}{|I_p - 2t_0 A\Sigma|^{\frac{1}{2}}}\Phi\left(\frac{\boldsymbol{\alpha}'(\Sigma^{-1} - 2t_0 A)^{-1}B'\mathbf{t}}{(1 + \boldsymbol{\alpha}'(\Sigma^{-1} - 2t_0 A)^{-1}\boldsymbol{\alpha})^{\frac{1}{2}}}\right), \end{aligned}$$

where $\mathbf{y} \sim N_p(\mathbf{0}, I)$. The last equality is obtained by using Lemma 1. □

From Theorem 1 we get the following three corollaries as special cases. First, substitute $\mathbf{t} = \mathbf{0}$ and we get the m.g.f. of the quadratic form $\mathbf{z}'A\mathbf{z}$.

Corollary 1.1. *Let $A : p \times p$ be a symmetric matrix and $\mathbf{z} \sim SN_p(\Sigma, \boldsymbol{\alpha})$. Then m.g.f. of $\mathbf{z}'A\mathbf{z}$ is*

$$M(t_0) = |I_p - 2t_0 A\Sigma|^{-\frac{1}{2}}. \tag{2.2}$$

If we consider $Z_1, \ldots, Z_p$ as random variables forming a sample of size p of dependent random variables, Corollary 1.1 will be useful to obtain the distribution of the sample variance of the sample $Z_1, \ldots, Z_p$ in the case when Z_i follow a skew normal distribution. We can write

$$S^2 = \frac{1}{p-1}\sum_{i=1}^{p}(Z_i - \bar{Z})^2 = \mathbf{z}'A\mathbf{z}$$

where $\bar{Z} = \frac{1}{p}\sum_1^p Z_i$ and

$$A = \frac{1}{p-1}\left(I_p - \frac{1}{p}\mathbb{I}_p\mathbb{I}_p'\right), \tag{2.3}$$

with I_p being the $p \times p$ identity matrix and $\mathbb{I}_p = (1, \ldots, 1)'$ the p-vector of ones.

When substituting $t_0 = 0$ in Theorem 1 we get the m.g.f. of the linear form $B\mathbf{z}$.

Corollary 1.2. *Let B be a $q \times p$-matrix and $\mathbf{z} \sim SN_p(\Sigma, \alpha)$. Then m.g.f. of $B\mathbf{z}$ is of the form*

$$M(\mathbf{t}) = 2e^{\frac{1}{2}\mathbf{t}'B\Sigma B'\mathbf{t}}\Phi\left(\frac{\alpha'\Sigma B'\mathbf{t}}{(1+\alpha'\Sigma\alpha)^{\frac{1}{2}}}\right).$$

It follows from (1.2) that $B\mathbf{z} \sim SN_q(B\Sigma B', \alpha^*)$ where $\alpha = B'\alpha^*$. In particular, if B: $p \times p$ is non-singular, then $B\mathbf{z} \sim SN_p(B\Sigma B', B'^{-1}\alpha)$. Furthermore, for $q = 1$, $B = \mathbf{h}'$, where $\mathbf{h}$ is a p-vector,

$$\mathbf{h}'\mathbf{z} \sim SN_1\left(\mathbf{h}'\Sigma\mathbf{h}, \alpha^*\right), \quad \text{where} \quad \alpha = \mathbf{h}'\alpha^*,$$

since

$$M_{\mathbf{h}'\mathbf{z}}(t) = 2e^{\frac{t^2}{2}\mathbf{h}'\Sigma\mathbf{h}}\Phi\left(\frac{\alpha'\Sigma\mathbf{h}}{(1+\alpha'\Sigma\alpha)^{\frac{1}{2}}}t\right), \quad t \in \mathbb{R}. \tag{2.4}$$

From (2.4) on can obtain the density of $\bar{Z} = \frac{1}{p}\sum_1^p Z_i$, where Z_i are not independent and $(Z_1, \ldots, Z_p)'$ follows a multivariate skew normal distribution. Here we can write $\bar{Z} = \frac{1}{p}\mathbb{I}_p'\mathbf{z}$, and $\bar{Z} \sim SN_1\left(\mathbb{I}_p'\Sigma\mathbb{I}_p, \alpha^*\right)$, where $\alpha = \mathbb{I}'\alpha^*$.

When $B = \mathbf{h}'$ and A is p.d. we get the next result.

Corollary 1.3. *Let $\mathbf{h}$ be a p-vector and A be a symmetric $p \times p$-matrix. If $\mathbf{z} \sim SN(\Sigma, \alpha)$ then the joint m.g.f. of $\mathbf{h}'\mathbf{z}$ and $\mathbf{z}'A\mathbf{z}$ is*

$$M(t_0, t_1) = \frac{2e^{\frac{t_1^2}{2}\mathbf{h}'(\Sigma^{-1}-2t_0A)^{-1}\mathbf{h}}}{|I_p - 2t_0A\Sigma|^{\frac{1}{2}}}\Phi\left(t_1\frac{\alpha'(\Sigma^{-1}-2t_0A)^{-1}\mathbf{h}}{(1+\alpha'(\Sigma^{-1}-2t_0A)^{-1}\alpha)^{\frac{1}{2}}}\right), \tag{2.5}$$

where $t_0, t_1 \in \mathbb{R}$.

From Theorem 1 we can conclude that the joint distribution of a linear and a quadratic forms in skew normal vector is uniquely determined by the m.g.f. (2.1). Namely, the probability measure is determined by the moment generating function if it exists in a neighborhood of 0 (Billingsley, 1986, p. 408). The last requirement is satisfied for the m.g.f. (2.1) as well as in special case (2.5). Another argument is used in Gupta, Nguyen, and Sanqui (2004) to show that skew normal distribution is uniquely determined by its sequence of moments. It follows from here that the joint distribution of the sample mean $\bar{Z} = \frac{1}{p}\mathbb{I}'_{\mathrm{p}}\mathbf{z}$ and the sample variance $DZ = \frac{1}{p-1}\sum_{i=1}^{p}(Z_i - \bar{Z})^2$ is also uniquely determined, i.e. when $\mathbf{h} = \frac{1}{p}\mathbb{I}_{\mathrm{p}}$ and A is given in (2.3).

3. Independence

Now we derive the conditions for the independence of a linear form and a quadratic form in a multivariate skew normal vector.

Theorem 2. *Let B be a $q \times p$-matrix and A be a symmetric $p \times p$-matrix. If $\mathbf{z} \sim SN_p(\Sigma, \alpha)$ then the linear form $B\mathbf{z}$ and the quadratic form $\mathbf{z}'A\mathbf{z}$ are independent if and only if $A\Sigma B' = \mathbf{0}$ and $B\Sigma\alpha = \mathbf{0}$ or $A\Sigma\alpha = \mathbf{0}$.*

Proof. For independent forms $B\mathbf{z}$ and $\mathbf{z}'A\mathbf{z}$ the joint m.g.f. is factorized into product of two terms

$$M(t_0, \mathbf{t}) = g_1(t_0)\, g_2(\mathbf{t}),$$

with g_1, g_2 being some functions of t_0 and $\mathbf{t}$ respectively, due to the definition of m.g.f. As the m.g.f. of the skew normal vector $\mathbf{z}$ determines uniquely the density function, also the opposite statement holds and independence of the forms under consideration follows from the factorization of m.g.f.

Let us examine the product $(\Sigma^{-1} - 2t_0A)^{-1}B'\mathbf{t}$ appearing in the m.g.f (2.1). We shall present the inverse $(\Sigma^{-1} - 2t_0A)^{-1}$ using the binomial inverse theorem (see, for example, Kollo and von Rosen, 2005, p. 75):

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(DA^{-1}B + C^{-1})^{-1}DA^{-1}, \tag{3.1}$$

where all the included inverses exist. When we take in (3.1) $A = \Sigma^{-1}$; $B = I_p$; $C = -2t_0I_p$; $D = A$ we get

$$(\Sigma^{-1} - 2t_0A)^{-1} = \Sigma - \Sigma(A\Sigma - (2t_0)^{-1}I_p)^{-1}A\Sigma = (I_p - (A - (2t_0)^{-1}\Sigma^{-1})^{-1}A)\Sigma. \tag{3.2}$$

When $A\Sigma B' = 0$, the m.g.f. (2.1) obtains the form

$$M(t_0, \mathbf{t}) = \frac{2e^{\mathbf{t}'B\Sigma B'\mathbf{t}}}{|I_p - 2t_0A\Sigma|^{\frac{1}{2}}}\Phi\left(\frac{\alpha'\Sigma B'\mathbf{t}}{(1 - \alpha'(\Sigma^{-1} - 2t_0A)^{-1}\alpha)^{\frac{1}{2}}}\right). \tag{3.3}$$

The right hand side of (3.3) can be factorized if and only if the argument of the distribution function does not depend on both variates t_0 and $\mathbf{t}$. When $B\Sigma\alpha = \mathbf{0}$, the argument equals zero and we have the m.g.f. of the form

$$M(t_0, \mathbf{t}) = \frac{e^{\mathbf{t}'B\Sigma B'\mathbf{t}}}{|I_p - 2t_0A\Sigma|^{\frac{1}{2}}}.$$

Another condition we get from the requirement that the argument does not depend on t_0. Equality (3.2) gives us the following expression in the numerator of argument of Φ:

$$1-\boldsymbol{\alpha}'(I_p-(A-(2t_0)^{-1}\Sigma^{-1})^{-1}A)\Sigma\boldsymbol{\alpha} = 1-\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha}+\boldsymbol{\alpha}'(A-(2t_0)^{-1}\Sigma^{-1})^{-1}A\Sigma\boldsymbol{\alpha}.$$

The obtained expression does not depend on t_0 if

$$A\Sigma\boldsymbol{\alpha} = \mathbf{0}.$$

In this case the m.g.f. is of the form

$$M(t_0,\mathbf{t}) = \frac{e^{\mathbf{t}'B\Sigma B'\mathbf{t}}}{|I_p - 2t_0 A\Sigma|^{\frac{1}{2}}}\Phi\left(\frac{\boldsymbol{\alpha}'\Sigma B'\mathbf{t}}{(1-\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha})^{\frac{1}{2}}}\right).$$

□

Remark 1. In the special case $B = \mathbf{h}'$ the dependence of a linear and a quadratic form in a skew normal vector is examined in Gupta and Huang (2002). We point out an erroneous statement in the paper. From the two conditions described above in Theorem 2, the requirement $\mathbf{h}'\Sigma\boldsymbol{\alpha} = \mathbf{0}$ is missing in Gupta and Huang (2002).

Corollary 2.1. *Let* $\mathbf{z} \sim SN_p(\Sigma,\boldsymbol{\alpha})$. *The linear form* $\mathbf{h}'\mathbf{z}$ *and the quadratic form* $\mathbf{z}'A\mathbf{z}$ *are independent if and only if* $A\Sigma\mathbf{h} = \mathbf{0}$ *and* $\mathbf{h}'\Sigma\boldsymbol{\alpha} = 0$ *or* $A\Sigma\boldsymbol{\alpha} = \mathbf{0}$.

4. Moments and cumulants

Moments and cumulants are found by differentiating the moment generating function (2.1) and cumulant generating function (c.g.f.) $K(t_0,\mathbf{t})$ of $B\mathbf{z}$ and $\mathbf{z}'A\mathbf{z}$, respectively . The c.g.f. is defined as

$$K(t_0,\mathbf{t}) = \ln M(t_0,\mathbf{t}).$$

Let us denote

$$\tilde{\mathbf{t}} = (t_0,\mathbf{t})';$$
$$W = \Sigma^{-1} - 2t_0 A$$

and

$$u = \frac{\boldsymbol{\alpha}'W^{-1}B'\mathbf{t}}{\sqrt{1+\boldsymbol{\alpha}'W^{-1}\boldsymbol{\alpha}}}.$$

As the moments can be presented through cumulants and vice versa (Kollo and von Rosen, 2005, p. 187), we shall differentiate the c.g.f. because of its simpler form. The joint c.g.f of $B\mathbf{z}$ and $\mathbf{z}'A\mathbf{z}$ is obtained from (2.1):

$$K(\tilde{\mathbf{t}}) = K(t_0,\mathbf{t}) = \ln 2+\frac{1}{2}\ln|W^{-1}\Sigma^{-1}|+\frac{1}{2}\mathrm{tr}(BW^{-1}B'\mathbf{t}\mathbf{t}')+\ln(\Phi(u)). \quad (4.1)$$

The two first cumulants are given in the next theorem.

Theorem 3. *Let* $\mathbf{z} \sim SN_p(\Sigma, \alpha)$*,* B *be a* $q \times p$*-matrix and* $A : p \times p$ *a p.d. matrix. Then the first two cumulants of the* $(q+1)$*-vector* $\mathbf{y} = (\mathbf{z}'A\mathbf{z}, B\mathbf{z})'$ *are:*

$$c_1(\mathbf{y}) = E\mathbf{y} = \begin{pmatrix} \operatorname{tr}(\Sigma A) \\ \sqrt{\frac{2}{\pi}} \frac{B\Sigma\alpha}{\sqrt{1+\alpha'\Sigma\alpha}} \end{pmatrix} \tag{4.2}$$

and

$$c_2(\mathbf{y}) = D\mathbf{y} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \Sigma_{22} \end{pmatrix} \tag{4.3}$$

where

$$\sigma_{11} = 2\operatorname{tr}(\Sigma A)^2;$$

$$\sigma_{12} = \sigma_{21}' = 2\sqrt{\frac{2}{\pi}} \left(\alpha'\Sigma A \Sigma B' - \frac{1}{2} \frac{\alpha'\Sigma A \Sigma \alpha \alpha' \Sigma B'}{1+\alpha'\Sigma\alpha} \right) \frac{1}{\sqrt{1+\alpha'\Sigma\alpha}};$$

$$\Sigma_{22} = B\Sigma B' - \frac{2}{\pi} \frac{B\Sigma\alpha\alpha'\Sigma B'}{1+\alpha'\Sigma\alpha}.$$

Proof. The derivatives of $K(\tilde{\mathbf{t}})$ are needed to find the cumulants. The two first derivatives of interest are organized as partitioned matrices

$$\frac{dK(\tilde{\mathbf{t}})}{d\tilde{\mathbf{t}}} = \begin{pmatrix} \dfrac{dK}{dt_0} \\ \dfrac{dK}{d\mathbf{t}} \end{pmatrix}$$

and

$$\frac{d^2K(\tilde{\mathbf{t}})}{d\tilde{\mathbf{t}}\, d\tilde{\mathbf{t}}'} = \begin{pmatrix} \dfrac{d^2K}{dt_0^2} & \dfrac{d^2K}{d\mathbf{t}\, dt_0} \\ \dfrac{d^2K}{dt_0\, d\mathbf{t}'} & \dfrac{d^2K}{d\mathbf{t}\, d\mathbf{t}'} \end{pmatrix}. \tag{4.4}$$

We will find the derivatives by arguments t_0 and $\mathbf{t}$ separately. For necessary properties of different matrix operations and derivatives the reader is referred to Kollo and von Rosen (2005) or Magnus and Neudecker (1999), for instance.

At the first step we find the first order derivative by the argument t_0 which corresponds to the quadratic form $\mathbf{z}'A\mathbf{z}$. Denoting

$$X = W^{-1}\Sigma^{-1} \tag{4.5}$$

and using the chain rule twice we get

$$\frac{dK(\tilde{\mathbf{t}})}{dt_0} = \frac{dX}{dt_0} \frac{dK(\tilde{\mathbf{t}})}{dX} \tag{4.6}$$

and

$$\frac{dX}{dt_0} = \frac{dX^{-1}}{dt_0}\frac{dX}{dX^{-1}}$$

where

$$\frac{dX^{-1}}{dt_0} = \frac{d(I_p - 2t_0\Sigma A)}{dt_0} = -2\text{vec}'(\Sigma A),$$

$$\frac{dX}{dX^{-1}} = -X \otimes X'.$$

Using the property of vec-operator and direct product

$$\text{vec}(ABC) = (C' \otimes A)\text{vec}(B),$$

we get

$$\frac{dX}{dt_0} = 2\text{vec}'(X\Sigma AX).$$

The last derivative in (4.6) equals

$$\begin{aligned}\frac{dK(\tilde{\mathbf{t}})}{dX} &= \frac{d}{dX}\left[\ln 2 + \frac{1}{2}\ln|X| + \frac{1}{2}\text{tr}(BX\Sigma B'\mathbf{t}\mathbf{t}') + \ln(\Phi(u)\right] \\ &= \frac{1}{2}\frac{d\ln|X|}{dX} + \frac{1}{2}\text{vec}(B'\mathbf{t}\mathbf{t}'B\Sigma) + \frac{du}{dX}\frac{\phi(u)}{\Phi(u)}\end{aligned} \tag{4.7}$$

where

$$\frac{d\ln|X|}{dX} = \frac{d|X|}{dX}\frac{1}{|X|} = |X|\text{vec}(X')^{-1}\frac{1}{|X|} = \text{vec}(X')^{-1}.$$

In the last term in (4.7) the derivative is

$$\begin{aligned}\frac{du}{dX} &= \frac{d}{dX}\left[\frac{\boldsymbol{\alpha}'X\Sigma B'\mathbf{t}}{\sqrt{1+\boldsymbol{\alpha}'X\Sigma\boldsymbol{\alpha}}}\right] \\ &= \frac{\Sigma B'\mathbf{t}\otimes\boldsymbol{\alpha}}{\sqrt{1+\boldsymbol{\alpha}'X\Sigma\boldsymbol{\alpha}}} - \frac{1}{2}\frac{\Sigma\boldsymbol{\alpha}\otimes\boldsymbol{\alpha}}{(1+\boldsymbol{\alpha}'X\Sigma\boldsymbol{\alpha})^{\frac{3}{2}}}\mathbf{t}'B\Sigma X'\boldsymbol{\alpha}.\end{aligned}$$

As a result the first order derivative of $K(\tilde{\mathbf{t}})$ by argument t_0 equals

$$\begin{aligned}\frac{dK(\tilde{\mathbf{t}})}{dt_0} &= 2\text{vec}'(X\Sigma AX)\left[\frac{1}{2}\text{vec}(X')^{-1} + \frac{1}{2}\text{vec}(B'\mathbf{t}\mathbf{t}'B\Sigma)\right. \\ &\left.+\left(\frac{\Sigma B'\mathbf{t}\otimes\boldsymbol{\alpha}}{\sqrt{1+\boldsymbol{\alpha}'X\Sigma\boldsymbol{\alpha}}} - \frac{1}{2}\frac{(\Sigma\boldsymbol{\alpha}\otimes\boldsymbol{\alpha})\mathbf{t}'B\Sigma X'\boldsymbol{\alpha}}{(1+\boldsymbol{\alpha}'X\Sigma\boldsymbol{\alpha})^{\frac{3}{2}}}\right)\frac{\phi(u)}{\Phi(u)}\right].\end{aligned} \tag{4.8}$$

The derivative by argument $\mathbf{t}$ corresponds to the linear form $B\mathbf{z}$. Using the matrix derivative of a product of matrices we get:

$$\begin{aligned}\frac{dK}{d\mathbf{t}} &= \frac{1}{2}\left((\mathbf{t}'\otimes I_q)+(I_q\otimes \mathbf{t}')\right)\mathrm{vec}(BW^{-1}B')+\frac{du}{d\mathbf{t}}\frac{\phi(u)}{\Phi(u)}\\ &= BW^{-1}B'\mathbf{t}+\frac{BW^{-1}\boldsymbol{\alpha}}{\sqrt{1+\boldsymbol{\alpha}'W^{-1}\boldsymbol{\alpha}}}\frac{\phi(u)}{\Phi(u)}.\end{aligned} \tag{4.9}$$

The first cumulant is obtained by substituting $\tilde{\mathbf{t}} = \mathbf{0}$ into (4.8) and (4.9):

$$c_1(\mathbf{y}) = \begin{pmatrix} \mathrm{tr}(\Sigma A) \\ \sqrt{\frac{2}{\pi}}\frac{B\Sigma\boldsymbol{\alpha}}{\sqrt{1+\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha}}}\end{pmatrix}.$$

The blocks in the second order derivative (4.4) are found in Appendix and presented by equalities (5.1), (5.3) and (5.5). To obtain the elements in the matrix of the second cumulant (4.3) these derivatives are evaluated at $\tilde{\mathbf{t}} = \mathbf{0}$: equalities (5.2), (5.4) and (5.6). □

From Theorem 3 we get the first two cumulants of the joint distribution of sample mean and variance for the sample $\mathbf{z} = (Z_1, \ldots, Z_p)'$ that follows the skew normal distribution $SN_p(\Sigma, \boldsymbol{\alpha})$.

Corollary 3.1. *Let $\mathbf{z} \sim SN_p(\Sigma, \boldsymbol{\alpha})$. Then for $\mathbf{y} = (S^2, \overline{Z})'$ we have the mean vector and the covariance matrix in the following form:*

$$E\mathbf{y} = E\begin{pmatrix} S^2 \\ \overline{Z}\end{pmatrix} = \begin{pmatrix} \frac{1}{p-1}\left(\mathrm{tr}\Sigma - \frac{1}{p}\mathrm{sum}(\Sigma)\right) \\ \frac{1}{p}\sqrt{\frac{2}{\pi}}\frac{\mathbb{I}_p'\Sigma\boldsymbol{\alpha}}{\sqrt{1+\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha}}}\end{pmatrix}, \quad \text{where} \quad \mathrm{sum}(\Sigma) = \sum_{i,j}\sigma_{ij}$$

and

$$D\mathbf{y} = D\begin{pmatrix} S^2 \\ \overline{Z}\end{pmatrix} = \begin{pmatrix} D(S^2) & \mathrm{cov}(S^2, \overline{Z}) \\ \mathrm{cov}(\overline{Z}, S^2) & D(\overline{Z})\end{pmatrix},$$

where $D(S^2)$ equals

$$D(S^2) = \frac{2}{(p-1)^2}\left(\mathrm{tr}(\Sigma^2) - \frac{1}{p}\mathrm{sum}(\Sigma^2)\right), \tag{4.10}$$

the second diagonal element $D(\overline{Z})$ equals

$$D(\overline{Z}) = \frac{1}{p^2}\left(\mathrm{sum}(\Sigma) - \frac{2}{\pi}\frac{(\mathbb{I}_p'\Sigma\boldsymbol{\alpha})^2}{1+\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha}}\right), \tag{4.11}$$

and the off-diagonal elements are:

$$\begin{aligned}\operatorname{cov}(S^2,\overline{Z}) &= \operatorname{cov}(\overline{Z},S^2)\\ &= \frac{2}{p(p-1)}\sqrt{\frac{2}{\pi(1+\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha})}}\\ &\times\left(\mathbb{I}_p'\Sigma^2\boldsymbol{\alpha}-\mathbb{I}_p'\Sigma\boldsymbol{\alpha}\left(\frac{1}{p}\operatorname{sum}(\Sigma)+\frac{1}{2}\frac{\boldsymbol{\alpha}'\Sigma^2\boldsymbol{\alpha}-\frac{1}{p}(\mathbb{I}_p'\Sigma\boldsymbol{\alpha})^2}{1+\boldsymbol{\alpha}'\Sigma\boldsymbol{\alpha}}\right)\right).\end{aligned} \tag{4.12}$$

Remark 2. In a special case when $B = 0$ the m.g.f. of $\mathbf{z}'A\mathbf{z}$ is given by (2.2). As the formula does not involve $\boldsymbol{\alpha}$, we have the same distribution as in the multivariate normal case. The cumulants $c_k(\mathbf{z}'A\mathbf{z})$ are thus known for any order $k = 1, 2, \dots$ (see, for example, Mathai and Provost, 1992, p. 52):

$$c_k(\mathbf{z}'A\mathbf{z}) = 2^{k-1}(k-1)!\ \operatorname{tr}(A\Sigma)^k.$$

This is in agreement with Genton et al. (2001), who derived

$$E(\mathbf{z}'A\mathbf{z}) = \operatorname{tr}(A\Sigma) \quad \text{and} \quad D(\mathbf{z}'A\mathbf{z}) = 2\operatorname{tr}(A\Sigma)^2.$$

Example. Let us consider a theoretical sample $\mathbf{z} \sim SN_n(\Sigma, \boldsymbol{\alpha})$, where $(\Sigma)_{ij} = (\theta^{|i-j|})$, $i, j = 1, \dots, n$. This way components of $\mathbf{z}$ form a finite autoregressive time series with parameter θ. We are interested in the joint density of the variance and the mean of that autoregressive time series $\mathbf{y} = (S^2, \overline{Z})'$.

Take the parameters $\theta = 0.6$ and $\boldsymbol{\alpha} = (7, \dots, 7)$, and the sample size $n = 10$.

Denote the correlation matrix

$$R\mathbf{y} = (D\mathbf{y})_d^{-1/2}(D\mathbf{y})(D\mathbf{y})_d^{-1/2},$$

where $(D\mathbf{y})_d$ is the diagonalized covariance matrix $D\mathbf{y}$. Using Corollary 3.1, the mean, covariance and correlation matrices of $\mathbf{y}$ are:

$$E\mathbf{y} = \begin{pmatrix}0.749\\0.455\end{pmatrix}, \quad D\mathbf{y} = \begin{pmatrix}0.210 & 0.003\\0.003 & 0.118\end{pmatrix}, \quad R\mathbf{y} = \begin{pmatrix}1 & 0.019\\0.019 & 1\end{pmatrix}.$$

In a simulation study we generate $k = 500\,000$ vectors of given size $n = 10$ from $SN_{10}(\Sigma, \boldsymbol{\alpha})$. Each generated vector plays a role of a sample with dependent elements. We shall calculate from each sample $\mathbf{z}_i$ the sample mean $\bar{z}$ and sample variance s_z^2. The obtained bivariate empirical distribution is given in the following Figure 1 by contour plots.

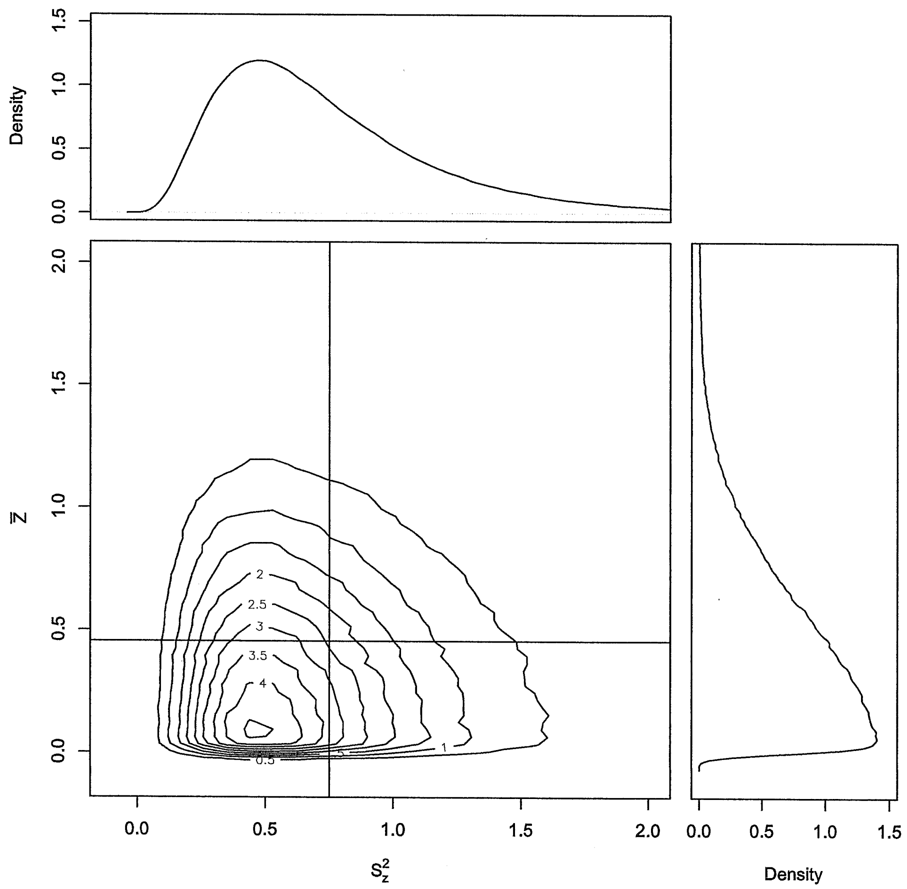

Figure 1. Empirical distribution of $\mathbf{y}$, with mean $E\mathbf{y}$ and marginal distributions.

As one can see, the empirical distribution is heavily skewed with marginal distributions of different type. For approximation of this distribution Edgeworth type approximations do not give a satisfactory solution as they are based on a multivariate distribution with marginals from the same class of distributions. From existing methods the copula theory can give an applicable model in this case (Nelsen, 1999).

5. Appendix: Second order derivatives of the joint c.g.f. of $B\mathbf{z}$ and $\mathbf{z}'A\mathbf{z}$

First we find second order derivative by the argument t_0 from c.g.f. $K(\tilde{\mathbf{t}})$ defined by (4.1). Matrix X, appearing in the further derivation, is given in

(4.5). The second order derivative by the argument t_0 is:

$$\begin{aligned}
\frac{d^2K(\tilde{\mathbf{t}})}{dt_0^2} &= \frac{dX}{dt_0}\frac{d}{dX}\left[\frac{dK(\tilde{\mathbf{t}})}{dt_0}\right] \\
&= \frac{dX}{dt_0}\frac{d}{dX}\Big[\mathrm{tr}(X'A\Sigma) + \mathrm{tr}(X'A\Sigma X'B'\mathbf{t}\mathbf{t}'B\Sigma) \\
&\quad +2\mathrm{vec}'(X\Sigma AX)\frac{du}{dX}\frac{\phi(u)}{\Phi(u)}\Big] \\
&= 2\mathrm{vec}'(X\Sigma AX)\Big[K_{p,p}\mathrm{vec}(\Sigma A) + \frac{d(X'A\Sigma X')}{dX}\mathrm{vec}(\Sigma B'\mathbf{t}\mathbf{t}'B) \\
&\quad + 2\frac{d}{dX}\left[\mathrm{vec}'(X\Sigma AX)\frac{du}{dX}\right]\frac{\phi(u)}{\Phi(u)} \qquad (5.1)\\
&\quad + 2\frac{du}{dX}\frac{\phi'(u)\Phi(u)-\phi^2(u)}{\Phi^2(u)}\left(\frac{du}{dX}\right)'\mathrm{vec}(X\Sigma AX)\Big] \\
&= 2\mathrm{vec}'(X\Sigma AX)\Big[\mathrm{vec}(A\Sigma) + \frac{d(X'A\Sigma X')}{dX}\mathrm{vec}(\Sigma B'\mathbf{t}\mathbf{t}'B) \\
&\quad + 2\left[\frac{d\mathrm{vec}'(X\Sigma AX)}{dX}\frac{du}{dX} + \frac{d^2u}{dX\,dX'}\mathrm{vec}(X\Sigma AX)\right]\frac{\phi(u)}{\Phi(u)} \\
&\quad + 2\frac{du}{dX}\frac{\phi'(u)\Phi(u)-\phi^2(u)}{\Phi^2(u)}\left(\frac{du}{dX}\right)'\mathrm{vec}(X\Sigma AX)\Big].
\end{aligned}$$

At the point $\tilde{\mathbf{t}} = \mathbf{0}$ we have

$$\left.\frac{d^2K(\tilde{\mathbf{t}})}{dt_0^2}\right|_{\tilde{\mathbf{t}}=\mathbf{0}} = 2\mathrm{tr}(\Sigma A)^2. \qquad (5.2)$$

The second order derivative by $\mathbf{t}$ equals:

$$\frac{d^2K(\tilde{\mathbf{t}})}{d\mathbf{t}\,d\mathbf{t}'} = BW^{-1}B' + \frac{BW^{-1}\alpha}{\sqrt{1+\alpha'W^{-1}\alpha}}\left(\frac{\phi'(u)}{\Phi(u)} - \frac{\phi^2(u)}{\Phi^2(u)}\right)\frac{\alpha'W^{-1}B'}{\sqrt{1+\alpha'W^{-1}\alpha}}. \qquad (5.3)$$

From here we get

$$\left.\frac{d^2K}{d\mathbf{t}\,d\mathbf{t}'}\right|_{\tilde{\mathbf{t}}=\mathbf{0}} = B\Sigma B' - \frac{2}{\pi}\frac{B\Sigma\alpha\alpha'\Sigma B'}{1+\alpha'\Sigma\alpha}. \qquad (5.4)$$

Finally we find the mixed derivative

$$\begin{aligned}\frac{d^2 K(\tilde{\mathbf{t}})}{d\mathbf{t}\, dt_0} &= \frac{dX}{dt_0}\frac{d}{dX}\left[BX\Sigma B'\mathbf{t} + \frac{BX\Sigma\alpha}{\sqrt{1+\alpha' X\Sigma\alpha}}\frac{\phi(u)}{\Phi(u)}\right] \\ &= 2\mathrm{vec}'(X\Sigma AX)\Bigg[\Sigma B'\mathbf{t}\otimes B' \\ &\quad + \left(\frac{\Sigma\alpha\otimes B'}{\sqrt{1+\alpha' X\Sigma\alpha}} - \frac{1}{2}\frac{(\Sigma\alpha\otimes\alpha)\,\alpha'\Sigma X' B'}{(1+\alpha' X\Sigma\alpha)^{\frac{3}{2}}}\right)\frac{\phi(u)}{\Phi(u)} \\ &\quad + \left(\frac{du}{dX}\frac{\phi'(u)\Phi(u)-\phi^2(u)}{\Phi^2(u)}\right)\frac{\alpha' X\Sigma B'}{\sqrt{1+\alpha' X\Sigma\alpha}}\Bigg],\end{aligned} \tag{5.5}$$

from where

$$\left.\frac{d^2 K}{d\mathbf{t}\, dt_0}\right|_{\tilde{\mathbf{t}}=\mathbf{0}} = 2\sqrt{\frac{2}{\pi}}\left(\alpha'\Sigma A\Sigma B' - \frac{1}{2}\frac{\alpha'\Sigma A\Sigma\alpha\alpha'\Sigma B'}{1+\alpha'\Sigma\alpha}\right)\frac{1}{\sqrt{1+\alpha'\Sigma\alpha}}. \tag{5.6}$$

Acknowlegements. The authors are thankful to the Referee whose comments helped to improve the presentation of the material. T. Kollo and A. Selart are grateful to the Estonian Science Foundation for financial support through grants ETF5686 and ETF7435.

References

Arellano-Valle, R. B., Branco, M. D., and Genton, M. G. (2006), *An unified view on skewed distributions arising from selections*, Canad. J. Statist. **34**, 581–601.

Arnold, B. C. and Beaver, R. J. (2002), *Skewed multivariate models related to hidden truncation and/or selective reporting*, Test **11**, 7–54.

Azzalini, A. and Dalla Valle, A. (1996), *The multivariate skew-normal distribution*, Biometrika **83**, 715–726.

Azzalini, A. and Capitanio, A. (1999), *Statistical applications of the multivariate skew normal distribution*, J. R. Stat. Soc. Ser. B **61**, 579–602.

Billingsley, P. (1986), *Probability and Measure*, Second Edition. Wiley & Sons, New York.

Genton, M. G., He, L., and Lin, X. (2001), *Moments of skew-normal random vectors*, Statist. Probab. Lett. **51**, 319–325.

Gupta, A. K. and Huang, W.-J. (2002), *Quadratic forms in multivariate skew normal variates*, J. Math. Anal. Appl. **273**, 558–564.

Gupta, A. K. and Kollo, T. (2003), *Density expansions based on multivariate skew normal distribution*, Sankhyā **65**, 821–835.

Gupta, A. K., Nguyen, T. T., and Sanqui, J. A. T. (2004), *Characterization of the skew-normal distribution*, Ann. Inst. Statist. Math. **56**, 351–360.

Kollo, T. and von Rosen, D. (2005), *Advanced Multivariate Statistics with Matrices.* Mathematics and Its Applications, 579, Springer, Dordrecht.

Knoth, S., Schmid, W., and Schöne, A. (2001), *Simultaneous Shewhart-type charts for the mean and the variance of a time series*, Frontiers Statist. Quality Control **6**, 61–79.

Magnus, J.R. and Neudecker, H. (1999), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Second edition. Wiley & Sons, Chichester.

Mathai, A. M. and Provost, S. B. (1992), *Quadratic Forms in Random Variables.* Marcel Dekker, New York

Nelsen, R. B. (1999), *An Introduction to Copulas.* Springer, New York.

Schöne, A. (1997), *Control charts for correlated data.* Ph.D. thesis, Department of Stochastics, University of Ulm.

Schöne, A. and Schmid, W. (2000), *On the joint distribution of a quadratic and a linear form in normal variables*, J. Multivariate Anal. **72**, 163–182.

Zacks, S. (1981), *Parametric Statistical Inference.* Pergamon Press, Oxford.

Department of Mathematics and Statistics, Bowling Green State University, USA

E-mail address: gupta@bgsu.edu

(T. Kollo and A. Selart) Institute of Mathematical Statistics, University of Tartu, J. Liivi 2, 50409 Tartu, Estonia

E-mail address: tonu.kollo@ut.ee

E-mail address: anne.selart@ut.ee