STUDIA MATHEMATICA 152 (2) (2002)

The Lukacs–Olkin–Rubin theorem without invariance of the "quotient"

by

Konstancja Bobecka and Jacek Wesołowski (Warszawa)

Abstract. The Lukacs theorem is one of the most brilliant results in the area of characterizations of probability distributions. First, because it gives a deep insight into the nature of independence properties of the gamma distribution; second, because it uses beautiful and non-trivial mathematics. Originally it was proved for probability distributions concentrated on $(0,\infty)$. In 1962 Olkin and Rubin extended it to matrix variate distributions. Since that time it has been believed that the fundamental reason such an extension is possible, is the assumed property of invariance of the distribution of the "quotient" (properly defined for matrices). The main result of this paper is that the matrix variate Lukacs theorem holds without any invariance assumption for the "quotient". The argument is based on solutions of some functional equations in matrix variate real functions, which seem to be of independent interest. The proofs use techniques of differential calculus in the cone of positive definite symmetric matrices.

1. Introduction. If for two independent, positive, nondegenerate random variables (rv's) their quotient and sum are also independent then the original rv's necessarily have the gamma distributions with the same scale parameter. This beautiful and important result, known as the Lukacs theorem, has drawn considerable attention of researchers in distribution theory since its publication in Lukacs (1955).

Further investigations of the characteristic property of the gamma law discovered by Lukacs were led at least in two basic directions: (i) weakening the independence condition to constancy of regressions—see Bolger and Harkness (1965), Hall and Simons (1969), Wesołowski (1990), Li, Huang, and Huang (1994), Huang and Su (1997) or Bobecka and Wesołowski (2001); (ii) considering rv's with values in more abstract structures as: $(0,\infty)^2$ in Wang (1981) or Bobecka (2001), stochastic processes in Wesołowski (1989), positive definite symmetric matrices in Olkin and Rubin (1962), symmetric cones in Casalis and Letac (1996) and Letac and Massam (1998).

2000 *Mathematics Subject Classification*: Primary 62H05.

Key words and phrases: Wishart distribution, positive definite symmetric matrices, independence, random matrices, matrix variate functions, functional equations.

In this paper we are concerned with matrix valued rv's.

Let $\mathcal{V}_+$ denote the cone of positive definite symmetric real $n\times n$ matrices. It is a subset of the Euclidean space $\mathcal{V}$ of symmetric real $n\times n$ matrices endowed with the inner product $(a,b)=\mathrm{trace}(ab)$ for $a,b\in\mathcal{V}$. Then the Lebesgue measure is introduced by assigning a unit mass to the unit cube in $\mathcal{V}$. Additionally denote by $\mathcal{M}$ the space of all $n\times n$ real matrices, which is Euclidean with the inner product $(a,b)=\mathrm{trace}(a^Tb)$, where ^{T} denotes transposition, and by $\overline{\mathcal{V}}_+$ the closed cone of positive symmetric $n\times n$ matrices. Let e be the identity matrix in $\mathcal{V}$. Define also $\mathcal{D}=\{x\in\mathcal{V}_+ :\ e-x\in\mathcal{V}_+\}$, which is an analogue of the interval $(0,1)$ in $\mathcal{V}$.

The Wishart (gamma) distribution $\gamma_{p,a}$ in $\overline{\mathcal{V}}_+$ is defined for any $a\in\mathcal{V}_+$ and any $p\in\Lambda_n=\{1/2,2/2,\ldots,(n-1)/2\}\cup((n-1)/2,\infty)$ by its Laplace transform as follows:

$$\int_{\overline{\mathcal{V}}_+}\exp(-(\theta,y))\,\gamma_{p,a}(dy)=(\det(e+\theta a^{-1}))^{-p},$$

for any $\theta+a\in\mathcal{V}_+$. If $p>(n-1)/2$ then $\gamma_{p,a}$ is absolutely continuous with respect to the Lebesgue measure and its density has the form

$$\gamma_{p,a}(dy)=\frac{(\det a)^p}{\Gamma_n(p)}(\det y)^{p-(n+1)/2}\exp(-(a,y))I_{\mathcal{V}_+}(y)\,dy,\qquad y\in\mathcal{V}_+,$$

where Γ_n is the n-variate Gamma function; see, for instance, Muirhead (1982), p. 61.

Olkin and Rubin (1962) formulated and proved a version of the Lukacs theorem in $\mathcal{V}_+$. As pointed out in Casalis and Letac (1996) their statement is not completely clear and the proof is even more difficult to follow. As these authors say, "... understanding the Olkin–Rubin proof (...) appears to be a strenuous task [we gave up after their identity (23)]". Instead they offered a clear statement of the Lukacs–Olkin–Rubin theorem together with a proof which is also not easy to follow unless one is familiar with analysis on symmetric cones and variance functions of natural exponential families (even the shorter proof given in Letac and Massam (1998) still needs considerable familiarity with these non-trivial domains). Here we state the Lukacs–Olkin–Rubin theorem following Casalis and Letac (1996):

Theorem 1. *Let X and Y be two independent rv's concentrated on $\overline{\mathcal{V}}_+$ such that $X+Y$ belongs to $\mathcal{V}_+$ almost surely and is not concentrated on some half line. Let $w:\mathcal{V}_+\to\mathcal{M}$ be a measurable function such that, for all $y\in\mathcal{V}_+$, one has $w(y)(w(y))^T=y$. Define*

$$U=(w(X+Y))^{-1}X((w(X+Y))^T)^{-1},\qquad V=X+Y.$$

If U and V are independent and are not concentrated on the same one-dimensional space and for any orthogonal matrix O the distributions of U

and OUO^T are the same then there exist $a \in \mathcal{V}_+$ and $p, q \in \Lambda_n$ such that $p+q > (n-1)/2$ and $X \sim \gamma_{p,a}$ and $Y \sim \gamma_{q,a}$.

The argument given in all the known proofs of this result, using the Laplace transform method, is heavily based on the invariance property of U, i.e. on the fact that U and OUO^T have the same distribution for any orthogonal $n \times n$ matrix O. Therefore for the last forty years there was a strong belief that this invariance property, which holds trivially in $(0,\infty)$, was rather fundamental for the Lukacs theorem in the case of matrix variate distributions, or more generally for distributions on symmetric cones. The main discovery of the present paper is that the invariance is not essential for characterizing the Wishart distribution.

Observe that one of the possible choices for w is $w(y) = y^{1/2}$ for $y \in \mathcal{V}_+$—possibly the most natural one. Our considerations will be restricted to this choice of w and to distributions having positive twice differentiable densities on $\mathcal{V}_+$. As mentioned above, typically problems related to the Lukacs theorem were attacked by Laplace transforms or characteristic functions. The idea of using the approach via densities came to our mind after successful use of smooth densities in proving a result, of a somewhat similar flavour to the Lukacs theorem, for the generalized inverse Gaussian and gamma laws, based on the so-called Matsumoto–Yor independence property —see Matsumoto and Yor (2001). It was accomplished essentially in Letac and Wesołowski (2000) and then refined, by considering a less restrictive smoothness assumption imposed on densities, in Wesołowski (2001). Let us point out that, additionally, the argument we offer here seems to be somewhat simpler, or at least more homogeneous, than the previous proofs of the original Lukacs–Olkin–Rubin theorem. It heavily depends on the methods of differential calculus in the cone of positive definite symmetric matrices. The statement and proof of our main result are given in Section 3. Section 2 is devoted to a thorough study of some functional equations on $\mathcal{V}_+$, which, while being of independent interest, play a crucial role in the proof of the strong version of the Lukacs–Olkin–Rubin theorem. Let us stress that the basic difficulty which has to be dealt with in the cone $\mathcal{V}_+$ is concerned with noncommutativity of multiplication, leading to non-trivial computations, while the case $n = 1$, i.e. $(0,\infty)$ is rather easy.

2. Functional equations. For any $y \in \mathcal{V}_+$ introduce linear operators $\mathbb{P}(y)$ and $\mathbb{L}(y)$ on $\mathcal{V}$, i.e. two elements of the space $L(\mathcal{V})$ of endomorphisms on $\mathcal{V}$, defined by

$$\mathbb{P}(y)h = yhy, \qquad \mathbb{L}(y)h = hy + yh,$$

for $h \in \mathcal{V}$. These two operators are important objects for the Jordan algebras theory and for analysis on symmetric cones of which the cone $\mathcal{V}_+$ is the most important example—see Faraut and Korányi (1994).

Observe that their inverses, denoted by $\mathbb{P}^{-1}(y)$ and $\mathbb{L}^{-1}(y)$, exist. We have $\mathbb{P}^{-1}(y) = \mathbb{P}(y^{-1})$ and the existence of $\mathbb{L}^{-1}(y)$ follows for instance from Theorem 5.1 of Olkin and Rubin (1964). Moreover, observe that the operators $\mathbb{P}$ and $\mathbb{L}$ are related by

$$\mathbb{P}(y^{-1}) \circ \mathbb{L}(y) = \mathbb{L}(y) \circ \mathbb{P}(y^{-1}) = \mathbb{L}(y^{-1}), \tag{2.1}$$

$$\mathbb{P}(y^{-1}) \circ \mathbb{L}(y^2) = \mathbb{L}(y^2) \circ \mathbb{P}(y^{-1}). \tag{2.2}$$

Also since $\mathbb{L}(y)$ is linear in y, we have $\mathrm{Id}_{\mathcal{V}} = \mathbb{L}(sy) \circ \mathbb{L}^{-1}(sy) = s\mathbb{L}(y) \circ \mathbb{L}^{-1}(sy)$ for any real s, where $\mathrm{Id}_{\mathcal{V}}$ is the identity in $L(\mathcal{V})$. Consequently,

$$\mathbb{L}^{-1}(sy) = s^{-1}\mathbb{L}^{-1}(y) \tag{2.3}$$

for any non-zero $s \in \mathbb{R}$ and any $y \in \mathcal{V}_+$.

Since $y = \mathbb{L}(y)(e/2)$, it follows that

$$\mathbb{L}^{-1}(y)y = e/2. \tag{2.4}$$

Finally observe that for any $y \in \mathcal{V}$,

$$\mathbb{L}^2(y) = \mathbb{L}(y^2) + 2\mathbb{P}(y). \tag{2.5}$$

These elementary properties of the operators $\mathbb{P}$ and $\mathbb{L}$ will be intensively exploited in the course of solution of two functional equations, which is the main objective of this section.

THEOREM 2. *Let $a : \mathcal{D} \to \mathbb{R}$ and $g : \mathcal{V}_+ \to \mathbb{R}$ be functions such that*

$$a(x) = g(yxy) - g(y(e-x)y) \tag{2.6}$$

for any $x \in \mathcal{D}$ and $y \in \mathcal{V}_+$. Assume that g is differentiable. Then there exist $\lambda, \beta \in \mathbb{R}$ such that for any $x \in \mathcal{D}$ and $y \in \mathcal{V}_+$,

$$a(x) = \lambda \log[\det x(e-x)^{-1}], \qquad g(y) = \lambda \log(\det y) + \beta.$$

Proof. Differentiation of (2.6) with respect to x (note that differentiability of a follows immediately from (2.6) since g is differentiable) gives

$$a'(x) = \mathbb{P}(y)[g'(\mathbb{P}(y)x) + g'(\mathbb{P}(y)(e-x))]. \tag{2.7}$$

Inserting $x = \frac{1}{2}e$ in (2.7) and replacing y by $(2y)^{1/2}$ we get

$$g'(y) = \mathbb{P}(y^{-1/2})b, \tag{2.8}$$

where $b = \frac{1}{4}a'(e/2) \in \mathcal{V}$.

We will show that there exists $\lambda \in \mathbb{R}$ such that $b = \lambda e$.

Observe that by taking $y = e$ in (2.7) we get

$$a'(x) = g'(x) + g'(e-x).$$

Inserting this identity back into (2.7) we obtain

$$g'(x) + g'(e-x) = \mathbb{P}(y)g'(\mathbb{P}(y)x) + \mathbb{P}(y)g'(\mathbb{P}(y)(e-x)),$$

which can be rewritten in the form

$$\mathbb{P}(y^{-1})g'(x) - g'(\mathbb{P}(y)x) = -[\mathbb{P}(y^{-1})g'(e-x) - g'(\mathbb{P}(y)(e-x))].$$

For any $s \in (0,1)$, change x to sx in the above equation. Observing that

$$g'(sx) = \mathbb{P}((sx)^{-1/2})b = s^{-1}\mathbb{P}(x^{-1/2})b = s^{-1}g'(x),$$

we see on multiplication by s that for any $s \in (0,1)$,

$$\mathbb{P}(y^{-1})g'(x) - g'(\mathbb{P}(y)x) = -s[\mathbb{P}(y^{-1})g'(e-sx) - g'(\mathbb{P}(y)(e-sx))].$$

Consequently, on letting $s \to 0$ and using the representation (2.8), we get

$$\mathbb{P}(y^{-1})g'(x) = g'(\mathbb{P}(y)x) \tag{2.9}$$

for any $x \in \mathcal{D}$, $y \in \mathcal{V}_+$.

Differentiate now (2.6) with respect to y to get

$$\begin{aligned} &xyg'(\mathbb{P}(y)x) + g'(\mathbb{P}(y)x)yx \\ &\qquad = (e-x)yg'(\mathbb{P}(y)(e-x)) + g'(\mathbb{P}(y)(e-x))y(e-x), \end{aligned}$$

which, due to (2.7) and (2.9), can be rewritten as

$$xy[\mathbb{P}(y^{-1})a'(x)] + [\mathbb{P}(y^{-1})a'(x)]yx = \mathbb{L}(y)\mathbb{P}(y^{-1})g'(e-x).$$

Hence by (2.1) we get

$$xa'(x)y^{-1} + y^{-1}a'(x)x = \mathbb{L}(y^{-1})g'(e-x).$$

And thus for any $y \in \mathcal{V}_+$,

$$y[xa'(x) - g'(e-x)] + [a'(x)x - g'(e-x)]y = 0.$$

Observe that if $yw + w^T y = 0$ for any $y \in \mathcal{V}_+$ and $w \in \mathcal{M}$ then $w = 0$. This follows on taking first $y = e$, which yields $w = -w^T$, and then $yw = wy$ for any $y \in \mathcal{V}_+$. Consequently (see the argument at the end of this proof), there exists $\lambda \in \mathbb{R}$ such that $w = \lambda e = w^T = -w = -\lambda e$ and hence $\lambda = 0$.

Therefore, since $a'(x)x - g'(e-x) = [xa'(x) - g'(e-x)]^T$ the above equation implies that

$$a'(x)x = g'(e-x) = xa'(x),$$

which further yields

$$xg'(e-x) = g'(e-x)x, \qquad x \in \mathcal{D}.$$

Thus, by substituting x for $e-x$ we have $(e-x)g'(x) = g'(x)(e-x)$. Hence $xg'(x) = g'(x)x$. Now (2.8) implies that $x(x^{-1/2}bx^{-1/2}) = (x^{-1/2}bx^{-1/2})x$, yielding $xb = bx$ for any $x \in \mathcal{D}$.

Take now for any $i, j \in \{1, \ldots, n\}$ the matrix $x_{i,j} = \frac{1}{2}e + \varepsilon e_{i,j}$, where $e_{i,j}$ is the matrix with all entries zero except the (i,j) and (j,i) entries which are one. Observe that $i = j$ is allowed and even necessary to consider in the case $n = 2$. Then for sufficiently small $\varepsilon > 0$ it follows that $x_{i,j} \in \mathcal{D}$ and substituting such x's in $bx = xb$, where $b = [b_{i,j}]$, gives $b_{i,k} = b_{j,k} = 0$ for any $k \neq i, j$ and $b_{i,i} = b_{j,j}$. Since this observation is valid for any i and j, it follows that b has to be a multiple of e.

Substituting $b = \lambda e$ where $\lambda \in \mathbb{R}$ in (2.8) gives $g'(y) = \lambda y^{-1}$. Hence

$$g(y) = \lambda \log(\det y) + \beta, \qquad \lambda, \beta \in \mathbb{R}. \tag{2.10}$$

Finally, inserting (2.10) into (2.6) we get

$$a(x) = \lambda \log[\det x(e-x)^{-1}]. \ \blacksquare$$

In the course of the proof of our next result we will need the following

PROPOSITION 1. *Assume that for some $b \in \mathcal{V}$,*

$$2\mathbb{P}(b)y = \mathbb{L}(b^2)y, \qquad y \in \mathcal{V}_+. \tag{2.11}$$

Then there exists $\lambda \in \mathbb{R}$ such that $b = \lambda e$.

Proof. By (2.11) we get, for any $y \in \mathcal{V}$,

$$\begin{aligned}
0 &= (\mathbb{L}(b^2)y - 2\mathbb{P}(b)y, y) = (b^2y, y) - (byb, y) + (yb^2, y) - (byb, y)\\
&= (yb, by) - (by, by) + (yb, by) - (yb, yb)\\
&= (yb - by, by) + (yb, by - yb) = (yb - by, by) - (yb - by, yb)\\
&= (yb - by, by - by) = (yb - by, (yb - by)^T) = \|yb - by\|^2,
\end{aligned}$$

where $\|\cdot\|$ denotes the norm defined by the inner product.

It follows that for any $y \in \mathcal{V}_+$ we have $by = yb$. Similarly to the final part of the preceding proof we conclude that $b = \lambda e$ for some real number λ. $\blacksquare$

Now we are ready to study the second functional equation; its proof is somewhat more involved than the previous one, though also based essentially on differentiation of functions of a matrix argument. But this time we will use second derivatives, i.e. instead of operating in the space $\mathcal{V}$ some of our equations in the proof will be in $L(\mathcal{V})$.

THEOREM 3. *Let $a_1 : \mathcal{D} \to \mathbb{R}$ and $a_2, g : \mathcal{V}_+ \to \mathbb{R}$ be functions satisfying*

$$a_1(x) + a_2(y) = g(yxy) + g(y(e-x)y) \tag{2.12}$$

for any $x \in \mathcal{D}$ and $y \in \mathcal{V}_+$. Assume that g is twice differentiable. Then there exist $\delta \in \mathcal{V}$ and $\lambda, \sigma, \sigma_1, \sigma_2 \in \mathbb{R}$ such that for any $x \in \mathcal{D}$ and $y \in \mathcal{V}_+$,

$$\begin{aligned}
a_1(x) &= \lambda \log[\det x(e-x)] + \sigma_1,\\
a_2(y) &= 4\lambda \log(\det y) + (\delta, y^2) + \sigma_2,\\
g(y) &= \lambda \log(\det y) + (\delta, y) + \sigma,
\end{aligned}$$

where $2\sigma = \sigma_1 + \sigma_2$.

Proof. Differentiation of (2.12) with respect to x gives

$$a_1'(x) = \mathbb{P}(y)[g'(\mathbb{P}(y)x) - g'(\mathbb{P}(y)(e-x))] \in \mathcal{V}. \tag{2.13}$$

Differentiating (2.13) once again with respect to x we arrive at

$$a_1''(x) = \mathbb{P}(y) \circ [g''(\mathbb{P}(y)x) + g''(\mathbb{P}(y)(e-x))] \circ \mathbb{P}(y) \in L(\mathcal{V}). \tag{2.14}$$

Substitute $x=\frac{1}{2}e$ in the above equation and then replace y by $(2y)^{1/2}$ to get

$$g''(y)=\mathbb{P}(y^{-1/2})\circ B\circ\mathbb{P}(y^{-1/2}), \tag{2.15}$$

where $B=\frac{1}{8}a_1''\big(\frac{1}{2}e\big)\in L(\mathcal{V})$.

Now our aim is to show that B is a multiple of $\mathrm{Id}_{\mathcal{V}}$, where $\mathrm{Id}_{\mathcal{V}}$ is the identity operator in $\mathcal{V}$.

Substituting $y=e$ in (2.14) gives

$$a_1''(x)=g''(x)+g''(e-x). \tag{2.16}$$

By inserting (2.16) back into (2.14) we obtain

$$g''(x)+g''(e-x)=\mathbb{P}(y)\circ[g''(\mathbb{P}(y)x)+g''(\mathbb{P}(y)(e-x))]\circ\mathbb{P}(y),$$

which can be rewritten in the form

$$\begin{aligned}g''(x)&-\mathbb{P}(y)\circ g''(\mathbb{P}(y)x)\circ\mathbb{P}(y)\\&=-[g''(e-x)-\mathbb{P}(y)\circ g''(\mathbb{P}(y)(e-x))\circ\mathbb{P}(y)].\end{aligned}$$

For any $s\in(0,1)$, change x to sx in the above equation. Since by (2.15),

$$g''(sx)=\mathbb{P}((sx)^{-1/2})\circ B\circ\mathbb{P}((sx)^{-1/2})=s^{-2}g''(x),$$

on multiplication by s^2 we deduce for any $s\in(0,1)$ that

$$\begin{aligned}g''(x)&-\mathbb{P}(y)\circ g''(\mathbb{P}(y)x)\circ\mathbb{P}(y)\\&=-s^2[g''(e-sx)-\mathbb{P}(y)\circ g''(\mathbb{P}(y)(e-sx))\circ\mathbb{P}(y)].\end{aligned}$$

Letting now $s\to 0$ and using (2.15) on the rhs, we obtain

$$\mathbb{P}(y^{-1})\circ g''(x)\circ\mathbb{P}(y^{-1})=g''(\mathbb{P}(y)x). \tag{2.17}$$

Now differentiate (2.13) with respect to y to get, for any $h\in\mathcal{V}$,

$$\begin{aligned}y[g'(\mathbb{P}(y)x)&-g'(\mathbb{P}(y)(e-x))]h+h[g'(\mathbb{P}(y)x)-g'(\mathbb{P}(y)(e-x))]y\\&=-\mathbb{P}(y)[g''(\mathbb{P}(y)x)+g''(\mathbb{P}(y)(e-x))][yxh+hxy]\\&\quad+\mathbb{P}(y)g''(\mathbb{P}(y)(e-x))[yh+hy].\end{aligned} \tag{2.18}$$

Observe now that by (2.13) and (2.14) equation (2.18) can be written as

$$\begin{aligned}y[\mathbb{P}(y^{-1})&a_1'(x)]h+h[\mathbb{P}(y^{-1})a_1'(x)]y\\&=-[a_1''(x)\mathbb{P}(y^{-1})][yxh+hxy]+\mathbb{P}(y)g''(\mathbb{P}(y)(e-x))\mathbb{L}(y)h.\end{aligned}$$

Then using (2.17) and (2.1) we get

$$a_1'(x)y^{-1}h+hy^{-1}a_1'(x)+a_1''(x)[y^{-1}hx+xhy^{-1}]=g''(e-x)\mathbb{L}(y^{-1})h.$$

Substitute $h=y$ in the above equation and use (2.15) to obtain

$$a_1'(x)+a_1''(x)x=\mathbb{P}((e-x)^{-1/2})B[(e-x)^{-1}].$$

Observe now that $a_1'(e-x) = -a_1'(x)$ and $a_1''(e-x) = a_1''(x)$. Then replacing x by $e-x$ above we arrive at

$$-a_1'(x) + a_1''(x)(e-x) = \mathbb{P}(x^{-1/2})B(x^{-1}).$$

Now using (2.16) and (2.15) we have

$$-a_1'(x) + [\mathbb{P}(x^{-1/2}) \circ B \circ \mathbb{P}(x^{-1/2}) + \mathbb{P}((e-x)^{-1/2}) \circ B \circ \mathbb{P}((e-x)^{-1/2})](e-x) \\ = \mathbb{P}(x^{-1/2})B(x^{-1}),$$

which implies

$$a_1'(x) = -\mathbb{P}(x^{-1/2})b + \mathbb{P}((e-x)^{-1/2})b, \tag{2.19}$$

where $b = B(e) \in \mathcal{V}$.

This equation will lead to a new formula for $a_1''(x)$ which will then be compared with the one obtained earlier. Observe first that

$$(x^{-1/2})'h = -\mathbb{P}(x^{-1/2})\mathbb{L}^{-1}(x^{1/2})h.$$

Now, using the above formula, differentiate (2.19) with respect to x and then compare to (2.14) to arrive via (2.15) at

$$\begin{aligned} x^{-1/2}b[\mathbb{P}(x^{-1/2})\mathbb{L}^{-1}(x^{1/2})h] + [\mathbb{P}(x^{-1/2})\mathbb{L}^{-1}(x^{1/2})h]bx^{-1/2} \\ + (e-x)^{-1/2}b[\mathbb{P}((e-x)^{-1/2})\mathbb{L}^{-1}((e-x)^{1/2})h] \\ + [\mathbb{P}((e-x)^{-1/2})\mathbb{L}^{-1}((e-x)^{1/2})h]b(e-x)^{-1/2} \\ = \mathbb{P}(x^{-1/2})B[\mathbb{P}(x^{-1/2})h] + \mathbb{P}((e-x)^{-1/2})B[\mathbb{P}((e-x)^{-1/2})h]. \end{aligned}$$

Insert now sx instead of x for $s \in (0,1)$ and use (2.3), then multiply both sides by s^2 and finally take $s \to 0$ to get, for any $x \in \mathcal{D}$,

$$bx^{-1/2}[\mathbb{L}^{-1}(x^{1/2})h] + [\mathbb{L}^{-1}(x^{1/2})h]x^{-1/2}b = B[\mathbb{P}(x^{-1/2})h]. \tag{2.20}$$

Replace first $x^{1/2}$ with x in (2.20), and then multiply both sides by t^{-2} for any positive number t. Then

$$b(tx)^{-1}[\mathbb{L}^{-1}(tx)h] + [\mathbb{L}^{-1}(tx)h](tx)^{-1}b = B[\mathbb{P}((tx)^{-1})h],$$

and, consequently,

$$by^{-1}[\mathbb{L}^{-1}(y)h] + [\mathbb{L}^{-1}(y)h]y^{-1}b = B[\mathbb{P}(y^{-1})h] \tag{2.21}$$

for any $y \in \mathcal{V}_+$. Substitute now $y = e$ in (2.21) to infer, since $\mathbb{L}^{-1}(e)h = h/2$, that

$$B(h) = \tfrac{1}{2}(bh + hb) = \tfrac{1}{2}\mathbb{L}(b)h \tag{2.22}$$

for any $h \in \mathcal{V}$. Evaluating now (2.21) at $h = \mathbb{L}(y)\mathbb{P}(y)b$ leads, via the commutativity property (2.1), to

$$2\mathbb{L}(b^2)y = \mathbb{L}(b)\mathbb{P}(y^{-1})\mathbb{L}(y)\mathbb{P}(y)b = \mathbb{L}(b)\mathbb{L}(y)b = \mathbb{L}(b^2)y + 2\mathbb{P}(b)y.$$

Consequently, $2\mathbb{P}(b)y = \mathbb{L}(b^2)y$ for all $y \in \mathcal{V}_+$ and by Proposition 1 it follows that $b = -\lambda e$ for some real number λ. Thus, (2.22) implies $B = -\lambda\, \mathrm{Id}_{\mathcal{V}}$.

Consequently, from (2.15) we get

$$g''(y) = \mathbb{P}(y^{-1/2})(-\lambda Id_{\mathcal{V}})\mathbb{P}(y^{-1/2}) = -\lambda\mathbb{P}(y^{-1}).$$

Observing that $(y^{-1})' = -\mathbb{P}(y^{-1})$, we get

$$g'(y) = \lambda y^{-1} + \delta, \qquad \delta \in \mathcal{V}.$$

Since $[\log(\det y)]' = y^{-1}$ and $[(\delta, y)]' = \delta$ we obtain

$$g(y) = \lambda \log(\det y) + (\delta, y) + \sigma, \qquad \sigma \in \mathbb{R}.$$

Then from (2.13) it follows that

$$a_1'(x) = \lambda[x^{-1} - (e-x)^{-1}],$$

and hence

$$a_1(x) = \lambda \log[\det x(e-x)] + \sigma_1, \qquad \sigma_1 \in \mathbb{R}.$$

Finally, from (2.12) we get

$$a_2(y) = 4\lambda \log(\det y) + (\delta, y^2) + \sigma_2,$$

where $\sigma_2 = 2\sigma - \sigma_1$. ∎

Remark 1. Observe that in the simplest case $n = 1$, i.e. $\mathcal{V}_+ = (0, \infty)$ and $\mathcal{D} = (0, 1)$, equations (2.6) and (2.12) can be written, respectively, as

$$a(x) = g(xy) - g((1-x)y), \qquad x \in (0,1),\ y > 0,$$

and

$$a_1(x) + a_2(y) = g(xy) + g((1-x)y), \qquad x \in (0,1),\ y > 0.$$

We are not aware of any result concerning solutions of such univariate equations, but with the differentiability assumption they can be solved fairly easily due to the fact that the multiplication is now commutative. These equations seem somewhat related to ones arising in problems connected with characterizations of information measures—see for instance Maksa (1987) or Kannappan and Sahoo (1993).

3. The matrix variate Lukacs theorem. The solutions to the functional equations, obtained in the previous section, will lead us to our main result, which is a new, strong version of the Lukacs–Olkin–Rubin theorem. Its basic feature, as mentioned in the introduction, is the lack of the assumption of the invariance of the distribution of the "quotient". This assumption was crucial in Olkin and Rubin (1962), Casalis and Letac (1996) or Letac and Massam (1998). On the other hand smoothness of densities is assumed. However, this assumption seems to be of rather technical nature, as the results related to the Matsumoto–Yor property suggest: the original assumption that the densities be continuously twice differentiable, imposed in Letac and Wesołowski (2000) has recently been reduced to just differentiability in Wesołowski (2001).

Also it is of considerable interest whether other division algorithms (see the formulation of Theorem 1), besides the one defined by $w(y) = y^{1/2}$ (the only one we are using here), can be dealt with by the approach through densities developed in this paper.

In what follows we will need the following result on the determinant of a certain linear mapping in $\mathcal{V}$; see Letac and Wesołowski (2000) for its proof.

PROPOSITION 2. *Let $c \in \mathcal{M}$. Denote by g_c the endomorphism of the linear space $\mathcal{V}$ defined by $x \mapsto cxc^T$. Then the absolute value of the determinant of g_c is $|\det c|^{n+1}$.*

Now we are ready to prove our main result.

THEOREM 4. *Let X and Y be independent rv's valued in $\mathcal{V}_+$ with strictly positive twice differentiable densities. Set $V = X + Y$ and $U = V^{-1/2}XV^{-1/2}$. If U and V are independent then there exist $p, q > (n-1)/2$ and $a \in \mathcal{V}_+$ such that $X \sim \gamma_{p,a}$ and $Y \sim \gamma_{q,a}$.*

Proof. Let $\psi : \mathcal{V}_+ \times \mathcal{V}_+ \to \mathcal{D} \times \mathcal{V}_+$ be defined by

$$\psi(x, y) = (u, v) = ((x + y)^{-1/2}x(x + y)^{-1/2}, x + y)$$

for $x, y \in \mathcal{V}_+$. Obviously $(U, V) = \psi(X, Y)$. Note that ψ is a bijection. Now our aim is to find the Jacobian of the map ψ^{-1}, which is defined by

$$(x, y) = (v^{1/2}uv^{1/2}, v^{1/2}(e - u)v^{1/2}) = (\mathbb{P}(v^{1/2})u, \mathbb{P}(v^{1/2})(e - u)),$$

that is, the determinant of the linear map

$$\begin{pmatrix} du \\ dv \end{pmatrix} \mapsto \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} dx/du & dx/dv \\ dy/du & dy/dv \end{pmatrix} \begin{pmatrix} du \\ dv \end{pmatrix}.$$

Since

$$\frac{dx}{du} = \mathbb{P}(v^{1/2}), \qquad \frac{dy}{du} = -\mathbb{P}(v^{1/2})$$

and

$$\frac{dy}{dv} = \mathrm{Id}_{\mathcal{V}} - \frac{dx}{dv},$$

we have

$$J = \begin{vmatrix} \mathbb{P}(v^{1/2}) & dx/dv \\ -\mathbb{P}(v^{1/2}) & \mathrm{Id}_{\mathcal{V}} - dx/dv \end{vmatrix} = \begin{vmatrix} \mathbb{P}(v^{1/2}) & dx/dv \\ 0 & \mathrm{Id}_{\mathcal{V}} \end{vmatrix} = \mathrm{Det}[\mathbb{P}(v^{1/2})],$$

where Det denotes the determinant in the space $L(\mathcal{V})$. Hence, from Proposition 2 we get

$$J = [\det v^{1/2}]^{n+1} = [\det v]^{(n+1)/2}.$$

Now we can find the joint density of (U, V). Since (X, Y) and (U, V) have independent components, the following identity holds for all $u \in \mathcal{D}$

and $v \in \mathcal{V}_+$:

$$f_U(u) f_V(v) = (\det v)^{(n+1)/2} f_X(v^{1/2} u v^{1/2}) f_Y(v^{1/2}(e-u)v^{1/2}), \tag{3.1}$$

where f_X, f_Y, f_U and f_V denote the densities of X, Y, U and V, respectively. Upon taking logarithms in (3.1) we get

$$g_1(u) + g_2(v) = g_3(\mathbb{P}(v^{1/2})u) + g_4(\mathbb{P}(v^{1/2})(e-u)), \tag{3.2}$$

where

$$g_1(u) = \log f_U(u), \tag{3.3}$$

$$g_2(v) = \log f_V(v) - \frac{n+1}{2} \log(\det v), \tag{3.4}$$

$$g_3 = \log f_X, \tag{3.5}$$

$$g_4 = \log f_Y. \tag{3.6}$$

Inserting $e-u$ for u in (3.2) gives

$$g_1(e-u) + g_2(v) = g_3(\mathbb{P}(v^{1/2})(e-u)) + g_4(\mathbb{P}(v^{1/2})u). \tag{3.7}$$

On subtracting (3.7) from (3.2) we obtain

$$\begin{aligned} & g_1(u) - g_1(e-u) \\ &= g_3(\mathbb{P}(v^{1/2})u) - g_4(\mathbb{P}(v^{1/2})u) - [g_3(\mathbb{P}(v^{1/2})(e-u)) - g_4(\mathbb{P}(v^{1/2})(e-u)]. \end{aligned} \tag{3.8}$$

Define

$$a(u) = g_1(u) - g_1(e-u), \qquad g = g_3 - g_4.$$

Then upon replacing v by $v^{1/2}$ equation (3.8) can be rewritten as

$$a(u) = g(\mathbb{P}(v)u) - g(\mathbb{P}(v)(e-u)).$$

Now, by Theorem 2 it follows that

$$a(u) = \lambda \log[\det u(e-u)^{-1}], \qquad g(v) = \lambda \log(\det v) + \beta,$$

for some $\lambda, \beta \in \mathbb{R}$. Hence

$$g_3(v) = g_4(v) + g(v) = g_4(v) + \lambda \log(\det v) + \beta. \tag{3.9}$$

Inserting (3.9) back into (3.2) gives

$$g_1(u) + g_2(v) = g_4(\mathbb{P}(v^{1/2})u) + \lambda \log[\det(\mathbb{P}(v^{1/2})u)] + \beta + g_4(\mathbb{P}(v^{1/2})(e-u)),$$

which can be rewritten in the form

$$a_1(u) + a_2(v) = g_4(\mathbb{P}(v^{1/2})u) + g_4(\mathbb{P}(v^{1/2})(e-u)), \tag{3.10}$$

where

$$a_1(u) = g_1(u) - \lambda \log(\det u), \tag{3.11}$$

$$a_2(v) = g_2(v) - \lambda \log(\det v) - \beta. \tag{3.12}$$

Replacing v by v^2 in (3.10) gives

$$a_1(u) + a_2(v^2) = g_4(\mathbb{P}(v)u) + g_4(\mathbb{P}(v)(e-u)).$$

Hence, by Theorem 3 it follows that

$$a_1(u) = \lambda_1 \log[\det u(e-u)] + \sigma_1,$$
$$a_2(v) = 2\lambda_1 \log(\det v) + (\delta, v) + \sigma_2,$$
$$g_4(v) = \lambda_1 \log(\det v) + (\delta, v) + \sigma,$$

for some $\delta \in \mathcal{V}$ and $\lambda_1, \sigma_1, \sigma_2 \in \mathbb{R}$, $\sigma_1 + \sigma_2 = 2\sigma$.

By (3.6) we get

$$\log f_Y(v) = \lambda_1 \log(\det v) + (\delta, v) + \sigma.$$

Hence

$$f_Y(v) = \det v^{\lambda_1} \exp(\delta, v) \exp \sigma,$$

and since f_Y is a density it follows that $\lambda_1 > -1$ and $a = -\delta \in \mathcal{V}_+$. Then $Y \sim \gamma_{q,a}$, where $q = \lambda_1 + (n+1)/2 > (n-1)/2$.

Now, since by (3.9),

$$g_3(v) = (\lambda + \lambda_1) \log(\det v) + (\delta, v) + \sigma + \beta,$$

from (3.5) it follows that

$$f_X(v) = [\det v]^{\lambda+\lambda_1} \exp(\delta, v) \exp(\sigma + \beta),$$

which implies that $\lambda + \lambda_1 > -1$ and consequently $X \sim \gamma_{p,a}$, while $p = \lambda + \lambda_1 + (n+1)/2 > (n-1)/2$. ■

REMARK 2. Recall that the matrix variate beta distribution $\beta_{p,q}$ on $\mathcal{D}$ is defined by the density

$$\beta_{p,q}(du) = \frac{(\det u)^{p-(n+1)/2}(\det(e-u))^{q-(n+1)/2}}{B_n(p,q)}\, du$$

where $p, q > (n-1)/2$ and $B_n(p,q)$ is the n-dimensional Euler beta function defined by

$$B_n(p,q) = \frac{\Gamma_n(p)\Gamma_n(q)}{\Gamma_n(p+q)}$$

(see for instance Muirhead (1981), p. 108).

Observe that in the above proof by (3.11) we have

$$g_1(u) = (\lambda + \lambda_1) \log[\det u] + \lambda_1 \log[\det(e-u)] + \sigma_1.$$

Then by (3.3),

$$f_U(u) = [\det u]^{\lambda+\lambda_1} [\det(e-u)]^{\lambda_1} \exp \sigma_1,$$

where $\lambda + \lambda_1 > -1$ and $\lambda_1 > -1$. Hence $U \sim \beta_{p,q}$ where $p = \lambda + \lambda_1 + (n+1)/2 > (n-1)/2$ and $q = \lambda_1 + (n+1)/2 > (n-1)/2$.

Also from (3.12) we get

$$g_2(v) = (\lambda + 2\lambda_1) \log(\det v) + (\delta, v) + \sigma_2 + \beta.$$

Hence by (3.4),

$$f_V(v) = [\det v]^{\lambda+2\lambda_1+(n+1)/2} \exp(\delta, v) \exp(\sigma_2 + \beta).$$

And thus $V \sim \gamma_{p+q,a}$, where as above $a = -\delta \in \mathcal{V}_+$.

References

K. Bobecka (2001), *Regression versions of Lukacs type characterizations for the bivariate gamma distribution*, J. Appl. Statist. Sci., to appear.

K. Bobecka and J. Wesołowski (2001), *Three dual regression schemes for the Lukacs theorem*, Metrika, to appear.

E. M. Bolger and W. L. Harkness (1965), *A characterization of some distributions by conditional moments*, Ann. Math. Statist. 36, 703–705.

M. Casalis and G. Letac (1996), *The Lukacs–Olkin–Rubin characterization of Wishart distributions on symmetric cones*, Ann. Statist. 24, 763–786.

J. Faraut and A. Korányi (1994), *Analysis on Symmetric Cones*, Oxford Univ. Press, New York.

W. J. Hall and G. Simons (1969), *On characterizations of the gamma distributions*, Sankhyā Ser. A 31, 385–390.

W.-J. Huang and J.-C. Su (1997), *On a study of renewal process connected with certain conditional moments*, ibid. 59, 28–41.

P. Kanappan and P. K. Sahoo (1993), *Sum form equations of multiplicative type*, Acta Math. Hungar. 61, 203–217.

G. Letac and H. Massam (1998), *Quadratic and inverse regressions for Wishart distributions*, Ann. Statist. 26, 573–595.

G. Letac and J. Wesołowski (2000), *An independence property for the product of GIG and gamma laws*, Ann. Probab. 28, 1371–1383.

S.-H. Li, W.-J. Huang and M.-N. L. Huang (1994), *Characterizations of the Poisson process as a renewal process via two conditional moments*, Ann. Inst. Statist. Math. 46, 351–360.

E. Lukacs (1955), *A characterization of the gamma distribution*, Ann. Math. Statist. 26, 319–324.

G. Maksa (1987), *The general solution of a functional equation arising in information theory*, Acta Math. Hungar. 49, 213–217.

H. Matsumoto and M. Yor (2001), *An analogue of Pitman's* $2M - X$ *theorem for exponential Wiener functionals. Part II*: *The role of the generalized inverse Gaussian laws*, Nagoya Math. J. 162, 65–86.

R. Muirhead (1982), *Aspects of Multivariate Analysis*, Wiley, New York.

I. Olkin and H. Rubin (1962), *A characterization of the Wishart distribution*, Ann. Math. Statist. 33, 1272–1280.

I. Olkin and H. Rubin (1964), *Multivariate beta distributions and independence properties of the Wishart distribution*, ibid. 35, 261–269.

Y. Wang (1981), *Extensions of Lukacs' characterization of the gamma distribution*, in: Analytic Methods in Probability Theory, Lecture Notes in Math. 861, Springer, New York, 166–177.

J. Wesołowski (1989), *A characterization of the gamma process by conditional moments*, Metrika 36, 299–309.

J. Wesołowski (1990), *A constant regression characterization of the gamma law*, Adv. Appl. Probab. 22, 488–490.

J. Wesołowski (2001), *The Matsumoto–Yor independence property for GIG and Gamma laws, revisited*, Math. Proc. Cambridge Philos. Soc., to appear.

Faculty of Mathematics and Information Science
Warsaw University of Technology
Pl. Politechniki 1
00-661 Warszawa, Poland
E-mail: wesolo@alpha.im.pw.edu.pl

Received May 4, 2001
Revised version January 2, 2002 (4733)