BU-313-M

CUMULANTS OF QUADRATIC AND BILINEAR

FORMS IN SINGULAR NORMAL VARIABLES[1]

S. R. Searle, N. S. Urquhart and D. A. Evans[2]
Biometrics Unit, Cornell University, Ithaca, N.Y. 14850

ABSTRACT

An expression is derived for the cumulant of a non-homogeneous quadratic form in normal variables that have a singular or non-singular variance-covariance matrix. Application to bilinear forms is shown, and the covariance between two bilinear forms is given.

- - - - - - - - - - -

1/ Paper number BU-205 in the Biometrics Unit and number 592 in the Department of Plant Breeding and Biometry, Cornell University.

2/ Present address: Dalgety Loan and Company, Wellington, New Zealand.

Also BU - 313-M

CUMULANTS OF QUADRATIC AND BILINEAR FORMS IN SINGULAR NORMAL VARIABLES[1/]

S. R. Searle, N. S. Urquhart and D. A. Evans[2/]
Biometrics Unit, Cornell University, Ithaca, N.Y. 14850

Summary

An expression is derived for the cumulant of a non-homogeneous quadratic form in normal variables that have a singular or non-singular variance-covariance matrix. Application to bilinear forms is shown, and the covariance between two bilinear forms is given.

\- - - - - - - - - - - -

The recent attention given to the singular normal distribution by Khatri (1963), Good (1963 and 1969), Rao (1966) and Shanbhag (1966 and 1969) centers on the independence of quadratic forms and on conditions under which quadratic forms have chi-squared distributions. In all of these papers, the normal distribution is assumed to have a mean vector of zero, a limitation not demanded by Rayner and Livingston (1965), who give conditions under which general non-homogeneous quadratic forms have χ^2-distributions. Although there is no conceptual loss of generality in assuming a mean of zero, because the nonzero case simply involves a translation, this translation alters both the distributional conditions and the cumulants.

1/ Paper number BU-205 in the Biometrics Unit and number 592 in the Department of Plant Breeding and Biometry, Cornell University.

2/ Present address: Dalgety Loan and Company, Wellington, New Zealand.

Rayner and Livingston (1965) consider the former but not the latter, which is what we deal with here. Starting from Lancaster (1954), who gives the cumulants of the quadratic form $\mathbf{x}'\mathbf{A}\mathbf{x}$ when $\mathbf{x}$ is normally distributed with a mean vector of zero and a variance-covariance matrix $\mathbf{I}$, we develop cumulants for the non-homogeneous quadratic form $\mathbf{x}'\mathbf{A}\mathbf{x} + \mathbf{b}'\mathbf{x} + c$ when $\mathbf{x}$ is a normally distributed random vector with a mean of $\boldsymbol{\mu}$, possibly non-null, and a covariance matrix $\mathbf{V}$, possibly singular. We write this as $\mathbf{x} \sim SN(\boldsymbol{\mu}, \mathbf{V})$, emphasizing the possible singularity of $\mathbf{V}$. Once Lancaster's results are so extended, they provide a basis for some generalizations and for obtaining some simple, but useful, results concerning bilinear forms.

1. Expected Values

Most of our discussion relates to $\mathbf{x}$ being a vector of normally distributed random variables. However, we first consider a result concerning the mean of a quadratic form which applies regardless of the underlying distribution.

Theorem 1 When $\mathbf{x}$ has mean $\boldsymbol{\mu}$ and variance-covariance matrix $\mathbf{V}$, i.e. when $\mathbf{x} \sim (\boldsymbol{\mu}, \mathbf{V})$,

$$E(\mathbf{x}'\mathbf{A}\mathbf{x}) = tr(\mathbf{V}\mathbf{A}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} . \tag{1}$$

Proof $$E(\mathbf{x}'\mathbf{A}\mathbf{x}) = Etr(\mathbf{x}\mathbf{x}'\mathbf{A}) = tr[E(\mathbf{x}\mathbf{x}')\mathbf{A}] = tr[(\mathbf{V} + \boldsymbol{\mu}\boldsymbol{\mu}')\mathbf{A}]$$

from which (1) follows at once.

Here, and in all that follows, considerable use is made of the trace of a matrix: $tr(\mathbf{X})$ equals the sum of the diagonal elements of $\mathbf{X}$. In particular we use the property of cyclic commutability of matrix products under the trace operation, *viz*, $tr(\mathbf{XYZ}) = tr(\mathbf{YZX}) = tr(\mathbf{ZXY})$; and also that a scalar is its own trace.

Corollary 1.1 When $\mu = 0$, i.e., when $x \sim (0, V)$, then $E(X'Ax) = tr(AV)$.

Corollary 1.2 When $\mu = 0$ and $V = I\sigma^2$, i.e. when $x \sim (0, I\sigma^2)$, then $E(x'Ax) = \sigma^2 tr(A)$ as given for example, in Graybill (1961, Theorem 4.18).

Corollary 1.3 The mean value of a non-homogeneous form is

$$E(x'Ax + b'x + c) = tr(AV) + \mu'A\mu + b'\mu + c .$$

2. Cumulants

In considering the quadratic form $x'Ax$, we assume without loss of generality that $A = A'$. Concerning the cumulants and their generating function we then have the following theorem.

Theorem 2 [Lancaster, (1954)] When $x \sim N(0, I)$, the cumulant generating function of $x'Ax$ is

$$K_{x'Ax}(t) = -\tfrac{1}{2}\log|I - 2itA| \tag{2}$$

and the s'th cumulant, the coefficient of $(it)^s/s!$ in (2), is

$$k_s(x'Ax) = 2^{s-1}(s - 1)!tr(A^s) . \tag{3}$$

Proof The proof is given in Lancaster (1954).

Corollary 2.1 When $x \sim SN(0, V)$ the parallel results are

$$K_{x'Ax}(t) = -\tfrac{1}{2}\log|I - 2itAV|, \text{ and } k_s(x'Ax) = 2^{s-1}(s - 1)!tr(AV)^s .$$

Proof Because V is positive semi-definite, $V = PP'$ for some matrix P of full column rank, the rank of V. Let $x = Py$ where $y \sim N(0, I)$ and apply Theorem 2.

Corollary 2.1 is a first, and simplest, generalization of Theorem 2. Successively wider generalization is achieved by considering $\underset{\sim}{x} \sim N(\underset{\sim}{\mu}, \underset{\sim}{I})$ then $\underset{\sim}{x} \sim SN(\underset{\sim}{\mu}, \underset{\sim}{V})$ and finally the non-homogeneous quadratic form $\underset{\sim}{x}'\underset{\approx}{A}\underset{\sim}{x} + \underset{\sim}{b}'\underset{\sim}{x} + c$. First, in the case of $\underset{\sim}{x} \sim N(\underset{\sim}{\mu}, \underset{\sim}{I}_m)$, for t sufficiently small the characteristic function of $\underset{\sim}{x}'\underset{\approx}{A}\underset{\sim}{x}$ is

$$\varphi_{\underset{\sim}{x}'\underset{\approx}{A}\underset{\sim}{x}}(t) = E(e^{it\underset{\sim}{x}'\underset{\approx}{A}\underset{\sim}{x}}) = (2\pi)^{-\frac{1}{2}m} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp[it\underset{\sim}{x}'\underset{\approx}{A}\underset{\sim}{x} - \tfrac{1}{2}(\underset{\sim}{x} - \underset{\sim}{\mu})'(\underset{\sim}{x} - \underset{\sim}{\mu})]dx_1 \cdots dx_m \, .$$

Writing the exponent as

$$\tfrac{1}{2}\underset{\sim}{\mu}'[(\underset{\sim}{I} - 2it\underset{\sim}{A})^{-1} - \underset{\sim}{I}]\underset{\sim}{\mu} - \tfrac{1}{2}[\underset{\sim}{x} - (\underset{\sim}{I} - 2it\underset{\sim}{A})^{-1}\underset{\sim}{\mu}]'(\underset{\sim}{I} - 2it\underset{\sim}{A})[\underset{\sim}{x} - (I - 2it\underset{\sim}{A})^{-1}\underset{\sim}{\mu}]$$

we see that the first part of it is constant with respect to integration over $\underset{\sim}{x}$ while the second part, after a translation, yields the same integrand as that associated with the characteristic function of $(\underset{\sim}{x} - \underset{\sim}{\mu})'\underset{\sim}{A}(\underset{\sim}{x} - \underset{\sim}{\mu})$. Thus the cumulant generating function of $\underset{\sim}{x}'\underset{\approx}{A}\underset{\sim}{x}$ is

$$K_{\underset{\sim}{x}'\underset{\approx}{A}\underset{\sim}{x}}(t) = \log \varphi_{\underset{\sim}{x}'\underset{\approx}{A}\underset{\sim}{x}}(t) = \tfrac{1}{2}\underset{\sim}{\mu}'[(\underset{\sim}{I} - 2it\underset{\sim}{A})^{-1} - \underset{\sim}{I}]\underset{\sim}{\mu} + K_{(\underset{\sim}{x}-\underset{\sim}{\mu})'\underset{\sim}{A}(\underset{\sim}{x}-\underset{\sim}{\mu})}(t) \, .$$

Now in general, for t small enough,

$$(\underset{\sim}{I} - 2it\underset{\sim}{A})^{-1} - \underset{\sim}{I} = \sum_{s=1}^{\infty} (2it)^s \underset{\sim}{A}^s \, , \qquad (4)$$

and $K_{(\underset{\sim}{x} - \underset{\sim}{\mu})'A(\underset{\sim}{x} - \underset{\sim}{\mu})}$ for $x \sim N(\underset{\sim}{\mu}, \underset{\sim}{I})$ is identical to (2). Therefore by identifying powers of it with the corresponding cumulants we have the s'th cumulant of $\underset{\sim}{x}'\underset{\approx}{A}\underset{\sim}{x}$ for $\underset{\sim}{x} \sim N(\underset{\sim}{\mu}, \underset{\sim}{I})$ as

$$k_s(\underset{\sim}{x}'\underset{\approx}{A}\underset{\sim}{x}) = 2^{s-1}(s - 1)![s\underset{\sim}{\mu}'\underset{\sim}{A}^s\underset{\sim}{\mu} + tr(\underset{\sim}{A}^s)] \, . \qquad (5)$$

Instead of next considering the cumulants of $\underset{\sim}{x}'\underset{\approx}{A}\underset{\sim}{x}$ for $\underset{\sim}{x} \sim SN(\underset{\sim}{\mu}, \underset{\sim}{V})$ we derive those of $Q \equiv \underset{\sim}{x}'\underset{\approx}{A}\underset{\sim}{x} + \underset{\sim}{b}'\underset{\sim}{x} + c$ for $\underset{\sim}{x} \sim N(\underset{\sim}{\mu}, \underset{\sim}{I})$ and then for $\underset{\sim}{x} \sim SN(\underset{\sim}{\mu}, \underset{\sim}{V})$. The cumulants of $\mathbf{x}'\mathbf{A}\mathbf{x}$ are then a special case of those of Q. Although Rayner and Livingstone (1965) discuss conditions under which Q has a χ^2-distribution, they do not consider the cumulants of Q. They can be derived by writing Q as the sum of a homogeneous quadratic form and a linear form that are independent. To achieve this we note, from Good (1963), that if $\mathbf{x} \sim N(\mathbf{0}, \mathbf{I})$, then $\mathbf{x}'\mathbf{M}\mathbf{x}$ and $\mathbf{p}'\mathbf{x}$ are independent if $\mathbf{M}\mathbf{p} = \mathbf{0}$, a result that also applies to the case when $\mathbf{x} \sim N(\boldsymbol{\mu}, \mathbf{I})$, as may be readily shown by a simple extension. With this in mind we write

$$
\begin{aligned}
Q &= \mathbf{x}'\mathbf{A}\mathbf{x} + \mathbf{b}'\mathbf{x} + c \\
&= (\mathbf{x} + \tfrac{1}{2}\mathbf{G}'\mathbf{b})'\mathbf{A}(\mathbf{x} + \tfrac{1}{2}\mathbf{G}'\mathbf{b}) + \mathbf{b}'(\mathbf{I} - \mathbf{G}\mathbf{A})\mathbf{x} + c - \tfrac{1}{4}\mathbf{b}'\mathbf{G}\mathbf{A}\mathbf{G}'\mathbf{b} \qquad (6)
\end{aligned}
$$

where $\mathbf{G}$ is a generalized inverse of $\mathbf{A}$,

$$\mathbf{A}\mathbf{G}\mathbf{A} = \mathbf{A} , \qquad (7)$$

and where, in this case, we also require that $\mathbf{G}$ satisfy

$$\mathbf{A}\mathbf{G}' = (\mathbf{G}\mathbf{A})' = \mathbf{G}\mathbf{A} . \qquad (8)$$

Such a $\mathbf{G}$ always exists for $\mathbf{A}$ being symmetric, as it is here [see Urquhart (1967)]. Then, because from the first and second terms of Q in (6),

$$\mathbf{A}[\mathbf{b}'(\mathbf{I} - \mathbf{G}\mathbf{A})]' = (\mathbf{A} - \mathbf{A}^2\mathbf{G}')\mathbf{b} = (\mathbf{A} - \mathbf{A}\mathbf{G}\mathbf{A})\mathbf{b} = \mathbf{0}$$

by (7) and (8), those two terms in Q are independent. Hence the cumulants of Q are the sums of the cumulants of its terms in (6); i.e.,

$$k_s(Q) = k_s([\mathbf{x} + \tfrac{1}{2}\mathbf{G}'\mathbf{b}]'\mathbf{A}[\mathbf{x} + \tfrac{1}{2}\mathbf{G}'\mathbf{b}]) + k_s(\mathbf{b}'[\mathbf{I} - \mathbf{G}\mathbf{A}]\mathbf{x}) + k_s(c - \tfrac{1}{4}\mathbf{b}'\mathbf{G}\mathbf{A}\mathbf{G}'\mathbf{b}) . \qquad (9)$$

Now with $\mathbf{x} \sim N(\boldsymbol{\mu}, \mathbf{I})$, it is obvious that $\mathbf{x} + \tfrac{1}{2}\mathbf{G}'\mathbf{b} \sim N(\boldsymbol{\mu} + \tfrac{1}{2}\mathbf{G}'\mathbf{b}, \mathbf{I})$, and so from (5)

$$k_s([\mathbf{x} + \tfrac{1}{2}\mathbf{G}'\mathbf{b}]'\mathbf{A}[\mathbf{x} + \tfrac{1}{2}\mathbf{G}'\mathbf{b}]) = 2^{s-1}(s - 1)![s(\boldsymbol{\mu} + \tfrac{1}{2}\mathbf{G}'\mathbf{b})'\mathbf{A}^s(\boldsymbol{\mu} + \tfrac{1}{2}\mathbf{G}'\mathbf{b}) + \mathrm{tr}(\mathbf{A}^s)] .$$

Also, because $\underline{x} \sim N(\underline{\mu}, \underline{I})$, (7) and (8) lead to

$$k_s(\underline{b}'[\underline{I} - \underline{GA}]\underline{x}) = \underline{b}'(\underline{I} - \underline{GA})\underline{\mu} \quad \text{for } s = 1 ;$$

$$= \underline{b}'(\underline{I} - \underline{GA})(\underline{I} - \underline{GA})'\underline{b} = \underline{b}'(\underline{I} - \underline{GA})\underline{b} \quad \text{for } s = 2 ;$$

$$= 0 \text{ for } s > 2 ;$$

and

$$k_s(c - \tfrac{1}{4}\underline{b}'\underline{GAGb}) = c - \tfrac{1}{4}\underline{b}'\underline{GAG}'\underline{b} \quad \text{for } s = 1 ;$$

$$= 0 \text{ for } s > 1 .$$

Substituting these values in (9) gives the following results:

$$E(Q) = k_1(Q)$$

$$= (\underline{\mu} + \tfrac{1}{2}\underline{G}'\underline{b})'\underline{A}(\underline{\mu} + \tfrac{1}{2}\underline{G}'\underline{b}) + tr(\underline{A}) + \underline{b}'(\underline{I} - \underline{GA})\underline{\mu} + c - \tfrac{1}{4}\underline{b}'\underline{GAG}'\underline{b}$$

$$= tr(\underline{A}) + \underline{\mu}'\underline{A\mu} + \underline{b}'\underline{\mu} + c , \tag{10}$$

in accord with Corollary 1.3 for $\underline{V} = \underline{I}$;

$$v(Q) = k_2(Q)$$

$$= 4(\underline{\mu} + \tfrac{1}{2}\underline{G}'\underline{b})'\underline{A}^2(\underline{\mu} + \tfrac{1}{2}\underline{G}'\underline{b}) + 2tr(\underline{A}^2) + \underline{b}'(\underline{I} - \underline{GA})\underline{b}$$

and on again using (7) and (8) this simplifies to

$$v(Q) = 2tr(\underline{A}^2) + 4(\tfrac{1}{2}\underline{b} + \underline{A\mu})'(\tfrac{1}{2}\underline{b} + \underline{A\mu}) ; \tag{11}$$

and for $s > 2$

$$k_s(Q) = 2^{s-1}(s - 1)![s(\underline{\mu} + \tfrac{1}{2}\underline{G}'\underline{b})'\underline{A}^s(\underline{\mu} + \tfrac{1}{2}\underline{G}'\underline{b}) + tr(\underline{A}^s)]$$

which, on using (7) and (8) again, reduces to

$$k_s(Q) = 2^{s-1}(s - 1)![s(\underline{\mu}'\underline{A}^s\underline{\mu} + \underline{b}'\underline{A}^{s-1}\underline{\mu} + \tfrac{1}{4}\underline{b}'\underline{A}^{s-2}\underline{b}) + tr(\underline{A})^s] .$$

$$= 2^{s-1}(s - 1)![s(\tfrac{1}{2}\underline{b} + \underline{A\mu})'\underline{A}^{s-2}(\tfrac{1}{2}\underline{b} + \underline{A\mu}) + tr(\underline{A})^s] , \tag{12}$$

with the usual convention that $\underset{\sim}{A}^0 = \underset{\sim}{I}$. Comparison of (11) and (12) shows that for s = 2, (12) reduces to (11). Hence (12) gives the s'th cumulant of Q for s > 1, and (10) gives the mean of Q. As a special case of (10) and (12) we have for $\underset{\sim}{\mu} = \underset{\sim}{0}$, i.e. for $\underset{\sim}{x} \sim N(\underset{\sim}{0}, \underset{\sim}{I})$

$$E(Q) = tr(\underset{\sim}{A}) + c$$

and (13)

$$k_s(Q) = 2^{s-1}(s-1)![\tfrac{1}{4}s\underset{\sim}{b}'\underset{\sim}{A}^{s-2}\underset{\sim}{b} + tr(\underset{\sim}{A})^s] \quad \text{for } s \geq 2 .$$

Now consider $\underset{\sim}{x} \sim SN(\underset{\sim}{\mu}, \underset{\sim}{V})$, with $\underset{\sim}{V} = \underset{\sim}{L}\underset{\sim}{L}'$ where $\underset{\sim}{L}$ is of full column rank (=m), the rank of $\underset{\sim}{V}$. Then, following Anderson (1958, p. 25) $\underset{\sim}{x}$ can be specified as

$$\underset{\sim}{x} = \underset{\sim}{\mu} + \underset{\sim}{L}\underset{\sim}{y} \text{ with } \underset{\sim}{y} \sim N(\underset{\sim}{0}, \underset{\sim}{I}_m) \ . \tag{14}$$

Note that $\underset{\sim}{x}$ cannot necessarily be specified as $\underset{\sim}{x} = \underset{\sim}{L}\underset{\sim}{y}$ as in the proof of Corollary 2.1. If it could, we would then have $\underset{\sim}{\mu} = E(\underset{\sim}{x}) = \underset{\sim}{L}E(\underset{\sim}{y}) = \underset{\sim}{L}\underset{\sim}{\theta}$, a relation that can fail if $\theta \not= 0$. For example if Z is a random variable distributed as N(0, 1), and $\underset{\sim}{x}' = (Z, Z + \alpha)$ for some constant α, then $\underset{\sim}{x} \sim SN(\underset{\sim}{\mu}, \underset{\sim}{V})$ where $\underset{\sim}{\mu}' = (0 \;\; \alpha)$ and $\underset{\sim}{V} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. But $\underset{\sim}{x}$ cannot be represented as $\underset{\sim}{x}' = \underset{\sim}{\ell}'Y$ not even for Y being some random variable distributed as $N(\theta, \sigma^2)$ for $E(\underset{\sim}{x}')$ would then be $\underset{\sim}{\ell}'\theta$ and no $\underset{\sim}{\ell}'$ exists such that $E(\underset{\sim}{x}') = \underset{\sim}{\ell}'\theta = (0 \;\; \alpha)$. But $\underset{\sim}{x}'$ <u>can</u> be represented as $\underset{\sim}{x}' = (0 \;\; \alpha) + (1 \;\; 1)Y$ for $Y \sim N(0 \;\; 1)$; i.e. as (14).

In terms of the specification of a singular normal distribution given in (14), the non-homogeneous quadratic of (6) is

$$\begin{aligned} Q &= \underset{\sim}{x}'\underset{\sim}{A}\underset{\sim}{x} + \underset{\sim}{b}'\underset{\sim}{x} + c \\ &= (\underset{\sim}{\mu} + \underset{\sim}{L}\underset{\sim}{y})'\underset{\sim}{A}(\underset{\sim}{\mu} + \underset{\sim}{L}\underset{\sim}{y}) + \underset{\sim}{b}'(\underset{\sim}{\mu} + \underset{\sim}{L}\underset{\sim}{y}) + c \\ &= \underset{\sim}{y}'\underset{\sim}{L}'\underset{\sim}{A}\underset{\sim}{L}\underset{\sim}{y} + (\underset{\sim}{b}' + 2\underset{\sim}{\mu}'\underset{\sim}{A})\underset{\sim}{L}\underset{\sim}{y} + \underset{\sim}{\mu}'\underset{\sim}{A}\underset{\sim}{\mu} + \underset{\sim}{b}'\underset{\sim}{\mu} + c \ . \end{aligned} \tag{15}$$

We now apply (13) to (15), using $\underline{L}'\underline{\underline{A}}\underline{L}$ for $\underline{\underline{A}}$ of (13), $(\underline{b}' + 2\underline{\mu}'\underline{\underline{A}})\underline{L}$ for $\underline{b}'$ of (13) and $\underline{\mu}'\underline{\underline{A}}\underline{\mu} + \underline{b}'\underline{\mu} + c$ for c of (13). This gives

$$E(Q) = \operatorname{tr}(\underline{L}'\underline{\underline{A}}\underline{L}) + \underline{\mu}'\underline{\underline{A}}\underline{\mu} + \underline{b}'\underline{\mu} + c$$

$$= \operatorname{tr}(\underline{\underline{V}}\underline{\underline{A}}) + \underline{\mu}'\underline{\underline{A}}\underline{\mu} + \underline{b}'\underline{\mu} + c \; ;$$

and for $s \geq 2$

$$k_s(Q) = 2^{s-1}(s-1)!\{s(\tfrac{1}{2}\underline{b} + \underline{\underline{A}}\underline{\mu})'\underline{L}(\underline{L}'\underline{\underline{A}}\underline{L})^{s-2}\underline{L}'(\tfrac{1}{2}\underline{b} + \underline{\underline{A}}\underline{\mu}) + \operatorname{tr}(\underline{L}'\underline{\underline{A}}\underline{L})^s\} \; . \tag{16}$$

Applying to (16) the results

$$\underline{L}(\underline{L}'\underline{\underline{A}}\underline{L})^r\underline{L}' = (\underline{\underline{V}}\underline{\underline{A}})^r\underline{\underline{V}} \text{ and } \operatorname{tr}(\underline{L}'\underline{\underline{A}}\underline{L})^r = \operatorname{tr}(\underline{\underline{A}}\underline{\underline{V}})^r$$

for positive integers r, leads to the following theorem.

Theorem 3 When $\underline{x} \sim SN(\underline{\mu}, \underline{\underline{V}})$, then $Q = \underline{x}'\underline{\underline{A}}\underline{x} + \underline{b}'\underline{x} + c$ has mean

$$E(Q) = \operatorname{tr}(\underline{\underline{V}}\underline{\underline{A}}) + \underline{\mu}'\underline{\underline{A}}\underline{\mu} + \underline{b}'\underline{\mu} + c$$

and cumulants, for integers $s \geq 2$,

$$k_s(Q) = 2^{s-1}(s-1)![s(\tfrac{1}{2}\underline{b} + \underline{\underline{A}}\underline{\mu})'(\underline{\underline{V}}\underline{\underline{A}})^{s-2}\underline{\underline{V}}(\tfrac{1}{2}\underline{b} + \underline{\underline{A}}\underline{\mu}) + \operatorname{tr}(\underline{\underline{V}}\underline{\underline{A}})^s] \; . \tag{17}$$

Here E(Q) is, of course, identical to the expression given in Theorem 1.

So far as cumulants of quadratic forms are concerned this is the most general result obtainable since it pertains to a non-homogeneous quadratic form in normal variables that have a mean which may be non-zero and a covariance matrix which may be singular. Manifold corollaries are available. We give but seven, listing, for the sake of completeness, some of the results given earlier.

Corollary 3.1 If $\underline{b} = \underline{A}\underline{g}$ and $c = \frac{1}{4}\underline{g}'\underline{A}\underline{g}$ for some $\underline{g}$, then $E(Q)$ and $k_s(Q)$ of (17) simplify to similar forms so that for integers $s \geq 1$

$$k_s(Q) = 2^{s-1}(s-1)![s(\tfrac{1}{2}\underline{g} + \underline{\mu})'\underline{A}(\underline{V}\underline{A})^{s-1}(\tfrac{1}{2}\underline{g} + \underline{\mu}) + tr(\underline{V}A)^s] .$$

Corollary 3.2 When $\underline{x} \sim SN(\underline{\mu}, \underline{V})$, the s'th cumulant of $\underline{x}'\underline{A}\underline{x}$ is

$$k_s(\underline{x}'\underline{A}\underline{x}) = 2^{s-1}(s-1)![s\underline{\mu}'\underline{A}(\underline{V}\underline{A})^{s-1}\underline{\mu} + tr(\underline{A}\underline{V})^s] ,$$

as derived in Rohde et al (1966).

Corollary 3.3 When $\underline{x} \sim N(\underline{\mu}, \underline{I})$ the s'th cumulant of $\underline{x}'\underline{A}\underline{x}$ is

$$k_s(\underline{x}'\underline{A}\underline{x}) = 2^{s-1}(s-1)![s\underline{\mu}'\underline{A}^s\underline{\mu} + tr(\underline{A})^s] .$$

This is the result developed in (5).

Corollary 3.4 When $\underline{x} \sim SN(\underline{0}, \underline{V})$

$$k_s(\underline{x}'\underline{A}\underline{x}) = 2^{s-1}(s - 1)!\,\mathrm{tr}(\underline{V}\underline{A})^s \ .$$

This is also corollary 2.1.

Corollary 3.5 When $\underline{x} \sim N(\underline{0}, \underline{I})$

$$k_s(\underline{x}'\underline{A}\underline{x}) = 2^{s-1}(s - 1)!\,\mathrm{tr}(\underline{A}^s) \ .$$

This is Theorem 2, given by Lancaster (1954)

Corollary 3.6 When $\underline{x} \sim SN(\underline{\mu}, \underline{V})$ the variance of $Q = \underline{x}'\underline{A}\underline{x} + \underline{b}'\underline{x} + c$ is

$$v(Q) = 2\mathrm{tr}(\underline{V}\underline{A})^2 + 4\underline{\mu}'\underline{A}\underline{V}\underline{A}\underline{\mu} + 4\underline{b}'\underline{V}\underline{A}\underline{\mu} + \underline{b}'\underline{V}\underline{b} \ .$$

Corollary 3.7 When $\underline{x} \sim SN(\underline{\mu}, \underline{V})$ the variance of $\underline{x}'\underline{A}\underline{x}$ is

$$v(\underline{x}'\underline{A}\underline{x}) = 2\mathrm{tr}(\underline{V}\underline{A})^2 + 4\underline{\mu}'\underline{A}\underline{V}\underline{A}\underline{\mu} \ .$$

The conditions under which $Q = \underline{x}'\underline{A}\underline{x} + \underline{b}'\underline{x} + c$ has a non-central χ^2-distribution can also be established from Theorem 3. The s'th cumulant of a non-central χ^2-distribution with q degrees of freedom and non-centrality parameter can be derived [from Graybill (1961, p. 76), for example] as

$$k_s = 2^{s-1}(s - 1)!(q + 2s\lambda) \ .$$

The necessary and sufficient conditions for this to equal, in Theorem 3, E(Q) for $s = 1$ and $k_s(Q)$ for $s > 1$ are exactly those under which Q has a non-central χ^2-distribution, as established by Rayner and Livingstone (1965, Theorem 7.2). These conditions, which are both necessary and sufficient, are

$$\begin{aligned}
&\text{(i)} \quad \underset{\sim}{V}\underset{\sim}{A}\underset{\sim}{V}\underset{\sim}{A}\underset{\sim}{V} = \underset{\sim}{V}\underset{\sim}{A}\underset{\sim}{V} \\
&\text{(ii)} \quad (\tfrac{1}{2}\underset{\sim}{b} + \underset{\sim}{A}\underset{\sim}{\mu})\underset{\sim}{V} = (\tfrac{1}{2}\underset{\sim}{b} + \underset{\sim}{A}\underset{\sim}{\mu})'\underset{\sim}{V}\underset{\sim}{A}\underset{\sim}{V} \\
\text{and} \qquad &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (18) \\
&\text{(iii)} \quad \underset{\sim}{\mu}'\underset{\sim}{A}\underset{\sim}{\mu} + \underset{\sim}{b}'\underset{\sim}{\mu} + c = (\tfrac{1}{2}\underset{\sim}{b} + \underset{\sim}{A}\underset{\sim}{\mu})'\underset{\sim}{V}(\tfrac{1}{2}\underset{\sim}{b} + \underset{\sim}{A}\underset{\sim}{\mu}) \ ,
\end{aligned}$$

the degrees of freedom of the resulting non-central χ^2-distribution being $tr(\underset{\sim}{V}\underset{\sim}{A})$, with non-centrality parameter $\frac{1}{2}(\underset{\sim}{\mu}'\underset{\sim}{A}\underset{\sim}{\mu} + \underset{\sim}{b}'\underset{\sim}{\mu} + c)$. The necessary and sufficient conditions for the χ^2-distribution to be a central χ^2 are the same as (18) except that (ii) and (iii) then have right-hand sides $\underset{\sim}{0}$ and 0 respectively. Unfortunately, Rao (1966a) has condition (ii) in this case as $\underset{\sim}{\mu}'\underset{\sim}{A}\underset{\sim}{V} = \underset{\sim}{0}$ when it should be $(\frac{1}{2}\underset{\sim}{b}' + \underset{\sim}{\mu}'\underset{\sim}{A})\underset{\sim}{V} = 0$, a correction that has been noted by Rao (1966b).

For the homogeneous quadratic form $\underset{\sim}{x}'\underset{\sim}{A}\underset{\sim}{x}$ the necessary and sufficient conditions of (18) reduce to

$$\underset{\sim}{V}\underset{\sim}{A}\underset{\sim}{V}\underset{\sim}{A}\underset{\sim}{V} = \underset{\sim}{V}\underset{\sim}{A}\underset{\sim}{V},\ \mu'\underset{\sim}{A}\underset{\sim}{V} = \mu'\underset{\sim}{A}\underset{\sim}{V}\underset{\sim}{A}\underset{\sim}{V} \text{ and } \underset{\sim}{\mu}'\underset{\sim}{A}\underset{\sim}{\mu} = \mu'\underset{\sim}{A}\underset{\sim}{V}\underset{\sim}{A}\underset{\sim}{\mu} \ . \qquad (19)$$

For non-singular $\underset{\sim}{V}$ these conditions reduce still further, to the equivalent conditions of idempotency of $\underset{\sim}{V}\underset{\sim}{A}$ or of $\underset{\sim}{V}$ being a generalized inverse of A, satisfying $\underset{\sim}{A}\underset{\sim}{V}\underset{\sim}{A} = \underset{\sim}{A}$. It is to be noted that this condition of idempotency, although both necessary and sufficient when $\underset{\sim}{V}$ is non-singular, is only sufficient when $\underset{\sim}{V}$ is singular; it is not necessary. [Equations (19) are then the necessary — and sufficient — conditions.] The error of describing it as a necessary condition made by Rao (1962) has been corrected by Rao (1966a).

4. Bilinear Forms

Quadratic forms are special cases of bilinear forms, and whereas the former have direct application in the estimation of variance components, for example, the latter can relate to estimation of covariance components. Furthermore, since a bilinear form in $\underline{x}'$ and $\underline{y}'$ can be expressed as a quadratic form in $(\underline{x}' \quad \underline{y}')$, properties of bilinear forms can readily be derived from those of quadratic forms, as we now illustrate.

We consider the general bilinear form $\underline{x}_1'A_{12}\underline{x}_2$ where $\underline{x}_1$ and $\underline{x}_2$ are vectors of order n_1 and n_2 respectively. Then $\underline{x}_1'\underline{A}_{12}\underline{x}_2$ can be expressed as a quadratic form

$$\underline{x}_1'\underline{A}_{12}\underline{x}_2 = \tfrac{1}{2}(\underline{x}_1' \quad \underline{x}_2')\begin{bmatrix} \underline{0} & \underline{A}_{12} \\ \underline{A}_{21} & \underline{0} \end{bmatrix}\begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} \text{ with } \underline{A}_{21} = \underline{A}_{12}' \ ;$$

i.e. as $\quad \underline{x}_1'\underline{A}_{12}\underline{x}_2 = \underline{x}'(\tfrac{1}{2}\underline{B})\underline{x}$

where

$$\underline{x}' = [\underline{x}_1' \quad \underline{x}_2'] \text{ and } \underline{B} = \underline{B}' = \begin{bmatrix} \underline{0} & \underline{A}_{12} \\ \underline{A}_{21} & \underline{0} \end{bmatrix} \text{ with } \underline{A}_{21} = \underline{A}_{12}' \ . \qquad (20)$$

Now suppose that

$$\underline{x} \sim SN(\underline{\mu}, \underline{V}) \text{ with } \underline{\mu} = \begin{bmatrix} \underline{\mu}_1 \\ \underline{\mu}_2 \end{bmatrix}, \ V = \begin{bmatrix} \underline{C}_{11} & \underline{C}_{12} \\ \underline{C}_{21} & \underline{C}_{22} \end{bmatrix} \text{ and } \underline{C}_{21} = \underline{C}_{12}' \ . \qquad (21)$$

This means that for i = 1 and 2, $E(\underline{x}_i) = \underline{\mu}_i$ and $E(\underline{x}_i - \underline{\mu}_i)(\underline{x}_i - \underline{\mu}_i)' = \underline{C}_{ii}$ and $E(\underline{x} - \underline{\mu}_1)(\underline{x}_2 - \underline{\mu}_2)' = \underline{C}_{12}$. Then properties of $\underline{x}_1'\underline{A}_{12}\underline{x}_{12}$ are equivalent to those of $\underline{x}'(\tfrac{1}{2}\underline{B})\underline{x}$, and can be derived from theorems 1 and 3 previously discussed. For example, from Theorem 1 we have the following.

Theorem 4 Whether the distribution of the x's is normal or not

$$E(\underline{x}_1'\underline{A}_{12}\underline{x}_2) = tr(\underline{A}_{12}\underline{C}_{21}) + \underline{\mu}_1'\underline{A}_{12}\underline{\mu}_2 \ .$$

Proof The proof follows directly from Theorem 1 using $\underline{B}$ and $\underline{V}$ of (20) and (21).

Theorem 3, with its wide span of generality, can provide many specific results concerning bilinear forms. We consider just one, from a corollary of which we then derive the covariance between 2 bilinear forms, that may or may not involve some of the same variables.

Theorem 5 The s-th cumulant of $\underline{x}_1'\underline{A}_{12}\underline{x}_2$ is

$$k_s(\underline{x}_1'\underline{A}_{12}\underline{x}_2) = \tfrac{1}{2}(s - 1)![s\underline{\mu}'\underline{B}(\underline{V}\underline{B})^{s-1}\underline{\mu} + \mathrm{tr}(\underline{V}\underline{B})^s]$$

with $\underline{B}$, $\underline{V}$ and $\underline{\mu}$ as defined in (20) and (21).

Proof Use $\tfrac{1}{2}\underline{B}$ for $\underline{A}$ in Corollary 3.2 .

Corollary 5.1 The variance of a bilinear form is

$$v(\underline{x}_1'\underline{A}_{12}\underline{x}_2) = \mathrm{tr}(\underline{A}_{12}\underline{C}_{21})^2 + \mathrm{tr}(\underline{A}_{12}\underline{C}_{22}\underline{A}_{21}\underline{C}_{11}) + \underline{\mu}_1'\underline{A}_{12}\underline{C}_{22}\underline{A}_{21}\underline{\mu}_1 + \underline{\mu}_2'\underline{A}_{21}\underline{C}_{11}\underline{A}_{12}\underline{\mu}_2 + 2\underline{\mu}_1'\underline{A}_{12}\underline{C}_{21}\underline{A}_{12}\underline{\mu}_2 \ . \qquad (22)$$

Proof This arises from substituting $\tfrac{1}{2}\underline{B}$ for $\underline{A}$ in Corollary 3.7 and using $(\underline{A}_{21})' = \underline{A}_{12}$, $(\underline{C}_{21})' = \underline{C}_{12}$ and the cyclic commutability of matrix products.

Theorem 5 and its corollary can be used in a variety of ways to obtain results for special cases. One instance of this is that it can yield the covariance between two bilinear forms, $\underline{x}_1'\underline{A}_{12}\underline{x}_2$ and $\underline{x}_3'\underline{A}_{34}\underline{x}_4$ say, for which we have the following theorem.

Theorem 6 Let $\underline{x} = \{\underline{x}_i\}$ and $\underline{\mu} = \{\underline{\mu}_i\}$ with $\underline{x}_i$ and $\underline{\mu}_i$ of order $n_i \times 1$ for i = 1,2,3 and 4; let $\underline{C} = \{\underline{C}_{ij}\}$ with $\underline{C}_{ij} = E(\underline{x}_i - \underline{\mu}_i)(\underline{x}_j - \underline{\mu}_j)' = (\underline{C}_{ji})'$ being $n_i \times n_j$, for i,j = 1,2,3 and 4. Then when $\underline{x} \sim SN(\underline{\mu}, \underline{C})$, the covariance between $\underline{x}_1'A_{12}\underline{x}_2$ and $\underline{x}_3\underline{A}_{34}\underline{x}_4$ is

$$\operatorname{cov}(\underline{x}_1'\underline{A}_{12}\underline{x}_2,\ \underline{x}_3'\underline{A}_{34}\underline{x}_4) = \operatorname{tr}(\underline{A}_{12}\underline{C}_{23}\underline{A}_{34}\underline{C}_{41} + \underline{A}_{12}\underline{C}_{24}\underline{A}_{43}\underline{C}_{31}) + \underline{\mu}_1'\underline{A}_{12}\underline{C}_{23}\underline{A}_{34}\underline{\mu}_4$$
$$+ \underline{\mu}_1'\underline{A}_{12}\underline{C}_{24}\underline{A}_{43}\underline{\mu}_3 + \underline{\mu}_2'\underline{A}_{21}\underline{C}_{13}\underline{A}_{34}\underline{\mu}_4 + \underline{\mu}_2'\underline{A}_{21}\underline{C}_{14}\underline{A}_{43}\underline{\mu}_3 \ . \tag{23}$$

Proof

$$\text{Let } \underline{W} = \frac{1}{2}\begin{bmatrix} \underline{0} & \underline{A}_{12} & \underline{0} & \underline{0} \\ (\underline{A}_{12})' & \underline{0} & \underline{0} & \underline{0} \\ \underline{0} & \underline{0} & \underline{0} & \underline{A}_{34} \\ \underline{0} & \underline{0} & (\underline{A}_{34})' & \underline{0} \end{bmatrix} .$$

Then

$$\underline{z}'\underline{W}\underline{z} = \underline{x}_1'\underline{A}_{12}\underline{x}_2 + \underline{x}_3'\underline{A}_{34}\underline{x}_4$$

so that

$$\operatorname{cov}(\underline{x}_1'\underline{A}_{12}\underline{x}_2,\ \underline{x}_3'\underline{A}_{34}\underline{x}_4) = \tfrac{1}{2}[v(\underline{z}'\underline{W}\underline{z}) - v(\underline{x}_1'\underline{A}_{12}\underline{x}_2) - v(\underline{x}_3'\underline{A}_{34}\underline{x}_4)] \ .$$

Applying Corollary 3.7 to $v(\underline{x}'\underline{W}\underline{x})$ and (22) to $v(\underline{x}_1'\underline{A}_{12}\underline{x}_2)$ and $v(\underline{x}_3'\underline{A}_{34}\underline{x}_4)$ yields, after a little simplification, (23).

We give only two of many possible corollaries to theorem 6:

Corollary 6.1

$$\operatorname{cov}(\underline{x}_1'\underline{A}\underline{x}_2,\ \underline{x}_1'\underline{B}\underline{x}_2) = \operatorname{tr}(\underline{A}\underline{C}_{21}\underline{B}\underline{C}_{21} + \underline{A}\underline{C}_{22}\underline{B}'\underline{C}_{11})$$
$$+ \underline{\mu}_1'\underline{A}\underline{C}_{21}\underline{B}\underline{\mu}_2 + \underline{\mu}_1'\underline{A}\underline{C}_{22}\underline{B}'\underline{\mu}_1$$
$$+ \underline{\mu}_2'\underline{A}'\underline{C}_{12}\underline{B}'\underline{\mu}_1 + \underline{\mu}_2'\underline{A}'\underline{C}_{11}\underline{B}\underline{\mu}_2 \ .$$

Corollary 6.2

$$v(\underline{x}'\underline{A}\underline{x}) = 2\operatorname{tr}(\underline{V}\underline{A})^2 + 4\underline{\mu}'\underline{A}\underline{V}\underline{A}\underline{\mu} \ ,$$

which is also corollary 3.7 .

Although other special cases could easily be considered, further explicit results for the general case seem of limited use. The results given here provide the means of dealing with special cases as they arise. In particular it is to be noted that the argument leading from Theorem 6 to Corollary 6.1, namely of using the vector $\underset{\sim}{x}' = (\underset{\sim}{x}_1' \ \underset{\sim}{x}_2' \ \underset{\sim}{x}_1' \ \underset{\sim}{x}_2')$, which has a singular covariance matrix, is one of the powerful features of the singular normal distribution: in this way it can be used to provide results for non-singular distributions.

References

Anderson, T. W. (1958). An Introduction to Multivariate Statistical Analysis. Wiley and Sons, New York.

Graybill, F. A. (1961). An Introduction to Linear Statistical Models. Vol. I. McGraw-Hill, New York.

Good, I. J. (1963). On the independence of quadratic expressions. J. Roy. Statist. Soc. Ser. B. 25, 377-382. Corrigenda 28 (1966), 584.

Good, I. J. (1969). Conditions for a quadratic form to have a chi-squared distribution. Biometrika, 56, 215-216.

Khatri, C. G. (1963). Further contributions to Wishartness and independence of second degree polynomials in normal vectors. J. Indian Stat. Ass. 1, 60-70.

Lancaster, H. O. (1954). Traces and cumulants of quadratic forms in normal variables. J. Roy. Statist. Soc. Ser. B. 16, 247-254.

Rayner, A. A. and D. Livingstone. (1965). On the distribution of quadratic forms in singular normal variables. S. African J. Agric. Sci., 8, 357-370.

Rohde, C. A., N. S. Urquhart and S. R. Searle. (1966). Quadratic forms in singular normal variables. Paper BU-261-M, Biometrics Unit Mimeo Series, Cornell University, Ithaca, New York.

Rao, C. R. (1962). A note on a generalized inverse of a matrix with applications to problems in mathematical statistics. J. Roy. Statist. Soc. Ser. B. 24, 152-58.

Rao, C. R. (1966a). Generalized inverse matrices and its applications in mathematical statistics in "Festschrift for J. Neyman". Ed. F. N. David, Wiley & Sons, New York, 263-279.

Rao, C. R. (1966b). Personal communication.

Shanbhag, D. N. (1966). On the independence of quadratic forms. _J. Roy. Statist. Soc. Ser. B_. 28, 582-583.

Shanbhag, D. N. (1968). Some remarks concerning Khatri's result on quadratic forms. _Biometrika_, 55, 593-595.

Urquhart, N. S. (1967). Computation of generalized inverse matrices which satisfy specified conditions. _Siam Review_, 10, 216-18.