The moments of products of quadratic forms in normal variables*

by JAN R. MAGNUS**

Abstract The expectation of the product of an arbitrary number of quadratic forms in normally distributed variables is derived.

1 Introduction

Tests of certain statistical hypotheses require a test-statistic which is a quadratic form. Typically, these tests arise in statistical inference on variances. For instance, let $\bar{x}$ and s^2 denote, respectively, the mean and the variance of a random sample $x_1, x_2, \ldots, x_n$ from an arbitrary distribution. Then,

$$ns^2 \equiv \sum_i (x_i - \bar{x})^2 = x'(\mathrm{I}_n - (1/n)1_n 1_n')x$$

is a quadratic form in x. The n-vector 1_n consists of ones only. If the sample arises from a normal distribution $N(\mu, \sigma^2)$, it is well-known that ns^2/σ^2 is distributed $\chi^2(n-1)$ regardless of the value of μ, a property very useful in the construction of confidence intervals for σ^2 when μ is not known.

As a second example, from econometrics, consider the linear regression model

$$y = X\beta + \varepsilon,$$

where X is a nonstochastic (n, k)-matrix of rank k and ε has a multivariate normal distribution $N_n(0, \sigma^2 V)$, while V is a positive definite known (n, n)-matrix. Let $\hat{\beta} = (X'V^{-1}X)^{-1}X'V^{-1}y$ be the GLS-estimator for β, and $e = y - X\hat{\beta}$ the vector of residuals. Then,

$$(n-k)\hat{\sigma}^2 \equiv e'V^{-1}e = \varepsilon'[V^{-1} - V^{-1}X(X'V^{-1}X)^{-1}X'V^{-1}]\varepsilon$$

is a quadratic form in ε. Moreover, the statistic $(n-k)\hat{\sigma}^2/\sigma^2$ is distributed $\chi^2(n-k)$.

Both examples are special cases of the following result, due to OGASAWARA and TAKAHASHI [6]:***

---

* Report AE4/78, Instituut voor Actuariaat en Econometrie, Universiteit van Amsterdam.

** Instituut voor Actuariaat en Econometrie, Universiteit van Amsterdam.

*** Even more general results are available, allowing for singular V. See RAO and MITRA [7, Theorem 9.2.1]. Note that condition (ii) in [7, p. 171]: "$V(A\mu + b)$ lies in the column space of VAV" can be replaced by the simpler expression: $V(A\mu + b) = VAV(A\mu + b)$.

Lemma 1.1

Let A be a symmetric (n, n)-matrix and $x \sim N_n(\mu, V)$, where V is positive definite (hence nonsingular). Necessary and sufficient that

$$x'Ax + 2b'x + \gamma$$

follows a $\chi^2(k, \delta)$ distribution is that

$$\begin{pmatrix} A \\ \cdots \\ b' \end{pmatrix} V(A : b) = \begin{pmatrix} A & b \\ b' & \gamma \end{pmatrix},$$

in which case

$$k = R(A), \text{ the rank of } A,$$

$$\delta = \mu' A \mu + 2b'\mu + \gamma.$$

Moreover, the distribution is central if and only if $A\mu = -b$.

The above two examples justify the study of quadratic forms in normal variables. In some cases the product of two or more quadratic forms is of interest. NAGAR [4] derived the expectation of $u'Au \cdot u'Bu$, A and B symmetric, and $u \sim N_n(0, I)$. NEUDECKER [5] found the expectation of $u'Au \cdot u'Bu \cdot u'Cu$ for arbitrary A, B, and C, while MAGNUS and NEUDECKER [3, theorem 4.2] simplified both proofs.

The purpose of the present paper is to derive, for arbitrary s, the expectation of $\prod_{j=1}^{s} \varepsilon' A_j \varepsilon$, where $\varepsilon \sim N_n(0, V)$ and $A_1 \ldots A_s$ are symmetric. Note that this formula can be used to compute all the moments of a product of an arbitrary number of quadratic forms. For example, in NAGAR's case, the variance of $u'Au \cdot u'Bu$ is easily determined from the formula.

Throughout the paper u denotes the standardnormal n-vector, $u \sim N_n(0, I_n)$, and ε will be distributed as $N_n(0, V)$, where V is a positive definite (hence nonsingular) matrix of order n. Further A, and $A_1, A_2, \ldots, A_s$ will denote real symmetric (n, n)-matrices, and Λ a real diagonal matrix. Of course, the symmetry assumption on A and $A_1 \ldots A_s$ is *not* restrictive in the study of quadratic forms.

The organization of this paper is as follows: In section two the moments of $\varepsilon' A \varepsilon$ are derived. In section three I introduce the concept of an $A(s)$-polynomial. The next section states a result from the theory of homogeneous symmetric functions in s variables. In section five the previous results are applied to prove the main result, i.e. to derive the expectation of $\prod_{j=1}^{s} \varepsilon' A_j \varepsilon$. The final section gives some applications.

2 The moments of $\varepsilon' A \varepsilon$

If the s-th moment of a random variable x exists, then its characteristic function $\varphi(t)$ can be expanded in a neighborhood of $t = 0$ as follows:

$$\varphi(t) = 1 + \sum_{h=1}^{s} \frac{(it)^h}{h!}\alpha_h + o(t^s), \tag{2.1}$$

where

$$\alpha_h = \mathrm{E}(x^h), \text{ the } h\text{-th moment of } x,$$

and

$$\lim_{t\to 0} \frac{o(t^s)}{t^s} = 0.$$

If $\varphi(t)$ can be expanded as in (2.1), then $\log\varphi(t)$ may be expanded as

$$\log\varphi(t) = \sum_{h=1}^{s} \frac{(it)^h}{h!}\kappa_h + o(t^s). \tag{2.2}$$

The quantities κ_h are called the *cumulants* (or *semi-invariants*) of the distribution of x. See CRAMÉR [1, p. 185–187]. Note that any cumulant κ_h is a polynomial in the moments $\alpha_1, \alpha_2, \ldots, \alpha_h$, and vice versa.

Lemma 2.1

The cumulants $\kappa_h(h = 1, 2, \ldots)$ of the distribution of $u'\Lambda u$, where $u \sim N_n(0, I)$ and Λ is diagonal, are

$$\boxed{\kappa_h(u'\Lambda u) = 2^{h-1}(h-1)!\,\mathrm{tr}\,\Lambda^h} \tag{2.3}$$

Proof

$$\begin{aligned}\kappa_h(u'\Lambda u) &= \kappa_h\Big(\sum_j \lambda_j u_j^2\Big) = \sum_j \kappa_h(\lambda_j u_j^2) = \sum_j \lambda_j^h \kappa_h(u_j^2)\\ &= \sum_j \lambda_j^h 2^{h-1}(h-1)! = 2^{h-1}(h-1)!\,\mathrm{tr}\,\Lambda^h.\end{aligned}$$

The second equality follows from the independence of the u_j^2 (CRAMÉR [1, p. 192]). The third equality also follows from [1, p. 187]. In the fourth equality I simply substitute the cumulants of the $\chi^2(1)$ distribution [1, p. 234].

Lemma 2.2

The s-th moment $\alpha_s = \mathrm{E}(u'\Lambda u)^s$ $(s = 1, 2, \ldots)$ is a known polynomial in $\kappa_1, \ldots, \kappa_s$ (see e.g. KENDALL and STUART [2, p. 69] for the first ten moments in terms of $\kappa_1 \ldots \kappa_{10}$). In particular we have

$$\alpha_1 = \mathrm{tr}\,\Lambda,$$

$$\alpha_2 = (\mathrm{tr}\,\Lambda)^2 + 2\mathrm{tr}\,\Lambda^2,$$

$$\alpha_3 = (\mathrm{tr}\,\Lambda)^3 + 6(\mathrm{tr}\,\Lambda)(\mathrm{tr}\,\Lambda^2) + 8\mathrm{tr}\,\Lambda^3,$$

$$\alpha_4 = (\mathrm{tr}\,\Lambda)^4 + 32(\mathrm{tr}\,\Lambda)(\mathrm{tr}\,\Lambda^3) + 12(\mathrm{tr}\,\Lambda^2)^2 + 12(\mathrm{tr}\,\Lambda)^2(\mathrm{tr}\,\Lambda^2) + 48\mathrm{tr}\,\Lambda^4.$$

Lemma 2.3

The s-th moment of $\varepsilon' A \varepsilon$, where $\varepsilon \sim N_n(0, V)$, is obtained from α_s (Lemma 2.2) by substituting AV for Λ. For example,

$$\mathrm{E}\varepsilon' A \varepsilon \quad = \operatorname{tr} AV,$$

$$\mathrm{E}(\varepsilon' A \varepsilon)^2 = (\operatorname{tr} AV)^2 + 2\operatorname{tr}(AV)^2, \text{ etc.}$$

Proof

Since V is positive definite, there exists a unique, positive definite and symmetric matrix $V^{\frac{1}{2}}$ such that $V^{\frac{1}{2}} V^{\frac{1}{2}} = V$. Let T be an orthogonal matrix such that

$$T' V^{\frac{1}{2}} A V^{\frac{1}{2}} T = \Lambda,$$

where Λ is a diagonal matrix containing the eigenvalues of $V^{\frac{1}{2}} A V^{\frac{1}{2}}$ on its diagonal. Then,

$$\varepsilon' A \varepsilon = (\varepsilon' V^{-\frac{1}{2}} T)(T' V^{\frac{1}{2}} A V^{\frac{1}{2}} T)(T' V^{-\frac{1}{2}} \varepsilon) = u' \Lambda u.$$

Hence,

$$\mathrm{E}(\varepsilon' A \varepsilon)^s = \mathrm{E}(u' \Lambda u)^s,$$

and

$$\operatorname{tr} \Lambda^h = \operatorname{tr}(T' V^{\frac{1}{2}} A V^{\frac{1}{2}} T)^h = \operatorname{tr}(AV)^h.$$

3 $A(s)$-forms and $A(s)$-polynomials

Consider now s real symmetric matrices $A_1, A_2, \ldots, A_s$. Before focusing on the expectation of $\prod_{j=1}^{s} (u' A_j u)$, where $u \sim N_n(0, I)$, I shall need some definitions that will prove useful in section five.

Definition 3.1 ($A(s)$-form)

Divide the index set $\{1, 2, \ldots, s\}$ into mutually exclusive and exhaustive subsets. Within each subset, take the trace of the matrix product of the A_h's corresponding with indices from this subset. The product of all these traces will be called an *$A(s)$-form.*

Examples of $A(3)$-forms:

$$\operatorname{tr}(A_1)\operatorname{tr}(A_2 A_3), \quad (\operatorname{tr} A_1)(\operatorname{tr} A_2)(\operatorname{tr} A_3), \quad \operatorname{tr}(A_1 A_2 A_3).$$

Definition 3.2 (similarity class)

Two $A(s)$-forms belong to the same similarity class iff their corresponding subsets (see definition 3.1) differ only by a permutation of indices.

Examples: $\operatorname{tr}(A_1)\operatorname{tr}(A_2 A_3)$ is equal to $\operatorname{tr}(A_3 A_2)\operatorname{tr} A_1$, but not necessarily equal to $\operatorname{tr}(A_2)\operatorname{tr}(A_1 A_3)$ or $\operatorname{tr}(A_3)\operatorname{tr}(A_1 A_2)$. However, all four $A(3)$-forms belong to the same similarity class. On the other hand, tr $(A_1A_2A_3)$ belongs to a different similarity class.

Definition 3.3 (*$A(s)$-sum*)

The sum of all non-equal $A(s)$-forms within a similarity class is called an *$A(s)$-sum*. Examples: The three $A(3)$-sums are: $\operatorname{tr}(A_1A_2A_3)$, $(\operatorname{tr} A_1)(\operatorname{tr} A_2)(\operatorname{tr} A_3)$, and $[(\operatorname{tr} A_1)\operatorname{tr}(A_2A_3)+(\operatorname{tr} A_2)\operatorname{tr}(A_1A_3)+(\operatorname{tr} A_3)\operatorname{tr}(A_1A_2)]$.

Definition 3.4 (*$A(s)$-polynomial*)

Any linear combination of $A(s)$-sums is called an *$A(s)$-polynomial.*
Examples: The $A(s)$-polynomials will play a crucial part in the remainder of this paper. Therefore I give the $A(2)$-, $A(3)$-, and $A(4)$-polynomials as examples:

$$A(2):\ v_1(\operatorname{tr} A_1)(\operatorname{tr} A_2)+v_2\operatorname{tr}(A_1A_2),$$

$$\begin{aligned}A(3):\ & v_1(\operatorname{tr} A_1)(\operatorname{tr} A_2)(\operatorname{tr} A_3)+v_2[(\operatorname{tr} A_1)(\operatorname{tr} A_2A_3)+(\operatorname{tr} A_2)(\operatorname{tr} A_1A_3)\\ & +(\operatorname{tr} A_3)(\operatorname{tr} A_1A_2)]+v_3\operatorname{tr}(A_1A_2A_3),\end{aligned}$$

$$\begin{aligned}A(4):\ & v_1(\operatorname{tr} A_1)(\operatorname{tr} A_2)(\operatorname{tr} A_3)(\operatorname{tr} A_4)+v_2[(\operatorname{tr} A_1)(\operatorname{tr} A_2A_3A_4)+(\operatorname{tr} A_2)(\operatorname{tr} A_1A_3A_4)\\ & +(\operatorname{tr} A_3)(\operatorname{tr} A_1A_2A_4)+(\operatorname{tr} A_4)(\operatorname{tr} A_1A_2A_3)]+v_3[(\operatorname{tr} A_1A_2)(\operatorname{tr} A_3A_4)\\ & +(\operatorname{tr} A_1A_3)(\operatorname{tr} A_2A_4)+(\operatorname{tr} A_1A_4)(\operatorname{tr} A_2A_3)]\\ & +v_4[(\operatorname{tr} A_1)(\operatorname{tr} A_2)(\operatorname{tr} A_3A_4)+(\operatorname{tr} A_1)(\operatorname{tr} A_3)(\operatorname{tr} A_2A_4)\\ & +(\operatorname{tr} A_1)(\operatorname{tr} A_4)(\operatorname{tr} A_2A_3)+(\operatorname{tr} A_2)(\operatorname{tr} A_3)(\operatorname{tr} A_1A_4)+(\operatorname{tr} A_2)(\operatorname{tr} A_4)(\operatorname{tr} A_1A_3)\\ & +(\operatorname{tr} A_3)(\operatorname{tr} A_4)(\operatorname{tr} A_1A_2)]+v_5[\operatorname{tr}(A_1A_2A_3A_4)+\operatorname{tr}(A_1A_2A_4A_3)+\operatorname{tr}(A_1A_3A_2A_4)].\end{aligned}$$

4 A Lemma

Before the theorem in the next section can be proved, I need one further result. It concerns a decomposition of a product of real variables into a polynomial of sums, and it may prove useful in other applications as well.

Let a, b, and c be real variables. It is easy to verify that

$$2ab=(a+b)^2-(a^2+b^2),$$

and

$$6abc=(a+b+c)^3-[(a+b)^3+(a+c)^3+(b+c)^3]+(a^3+b^3+c^3).$$

This suggests the following

Lemma 4.1

Let $x_1, x_2, \ldots$ be real variables. Then, for $s=2, 3, \ldots$

$$\boxed{s!\prod_{j=1}^{s} x_j = \sum_{k=0}^{s-1}\left\{(-1)^k \sum_{i_1<i_2<\ldots<i_k}^{s}\left(\sum_{\substack{j=1\\ j\neq i_1\neq\ldots\neq i_k}}^{s} x_j\right)^{s}\right\}} \tag{4.1}$$

Proof

Denote the r.h.s. of (4.1) as $\varphi(x_1, x_2, \ldots, x_s)$. I wish to prove that $\varphi(x_1, x_2, \ldots, x_s) = s!\prod_{j=1}^{s} x_j$. Suppose that $x_s = 0$ implies $\varphi(x_1, x_2, \ldots, x_s) = 0$. Then, by symmetry, the same is true for $x_1, x_2, \ldots, x_{s-1}$, and hence

$$\varphi(x_1, x_2, \ldots, x_s) = C\prod_{j=1}^{s} x_j,$$

where C is a constant. Now, the only term in $\varphi(x_1, x_2, \ldots, x_s)$ which yields $\Pi_j x_j$ is $(\Sigma_j x_j)^s$. Hence, the coefficient of $\Pi_j x_j$ must be $s!$. The only thing to be shown, then, is that $x_s = 0$ implies $\varphi(x_1, x_2, \ldots, x_s) = 0$.

Define the function $\varphi_k^{(\tau)}(\cdot)$ in $s-1$ variables as

$$\varphi_k^{(\tau)}(x_1, x_2, \ldots, x_{s-1}) = \sum_{i_1<i_2<\ldots<i_k}^{\tau}\left(\sum_{\substack{j=1\\ j\neq i_1\neq\ldots\neq i_k}}^{s-1} x_j\right)^{s}, \qquad \begin{array}{l} k = 0, 1, \ldots, s-1; \\ \tau = s-1, s. \end{array} \tag{4.2}$$

The following properties of $\varphi_k^{(\tau)}(\cdot)$ are readily verified:

$$\varphi_k^{(s)}(\cdot) = \varphi_k^{(s-1)}(\cdot) + \varphi_{k-1}^{(s-1)}(\cdot), \qquad k = 1, 2, \ldots, s-2, \tag{4.3}$$

$$\varphi_0^{(s)}(\cdot) = \varphi_0^{(s-1)}(\cdot) = \left(\sum_{j=1}^{s-1} x_j\right)^{s}, \tag{4.4}$$

$$\varphi_{s-1}^{(s)}(\cdot) = \varphi_{s-2}^{(s-1)}(\cdot) = \sum_{j=1}^{s-1} x_j^{s}. \tag{4.5}$$

Hence, using the definition (4.2) and the properties (4.3), (4.4), and (4.5),

$$\begin{aligned}
\varphi(x_1, x_2, \ldots, x_{s-1}, 0) &= \sum_{k=0}^{s-1}(-1)^k \varphi_k^{(s)}(x_1, x_2, \ldots, x_{s-1}) \\
&= \varphi_0^{(s)}(\cdot) + \sum_{k=1}^{s-2}(-1)^k[\varphi_k^{(s-1)}(\cdot) + \varphi_{k-1}^{(s-1)}(\cdot)] + (-1)^{s-1}\varphi_{s-1}^{(s)}(\cdot) \\
&= \varphi_0^{(s)}(\cdot) - \varphi_0^{(s-1)}(\cdot) + (-1)^{s-2}\varphi_{s-2}^{(s-1)}(\cdot) + (-1)^{s-1}\varphi_{s-1}^{(s)}(\cdot) = 0.
\end{aligned}$$

This concludes the proof.

5 The expectation of $\prod_{j=1}^{s} \varepsilon' A_j \varepsilon$

The main result then is the following:

Theorem 5.1

Let $A_1, A_2, \ldots, A_s$ be real symmetric (n, n)-matrices, $u \sim N_n(0, I)$, and $\varepsilon \sim N_n(0, V)$, V positive definite. Then,

$$\mathrm{E} \prod_{j=1}^{s} u'A_j u \qquad \text{is an } A(s)\text{-polynomial,}$$

and

$$\mathrm{E} \prod_{j=1}^{s} \varepsilon' A_j \varepsilon \qquad \text{is the same } A(s)\text{-polynomial with each } A_i \text{ replaced}$$

by $A_i V$.

Proof

Put $x_j = u'A_j u$ in lemma 4.1 and take expectations. This gives

$$s!\,\mathrm{E} \prod_{j=1}^{s} (u'A_j u) = \sum_{k=0}^{s-1} \left\{ (-1)^k \sum_{i_1 < i_2 < \ldots < i_k}^{s} \mathrm{E} \left[u' \left(\sum_{\substack{j=1 \\ j \neq i_1 \neq \ldots \neq i_k}}^{s} A_j \right) u \right]^s \right\}. \tag{5.1}$$

Note that (5.1) holds generally, not only for normally distributed u. Also note that (5.1) in conjunction with lemma 2.3 shows that the expectation of $\prod_{j=1}^{s} u'A_j u$ is a sum of terms, each of which is either an $A(s)$-form (see definition 3.1) or has the same structure as an $A(s)$-form but differs from it in that at least one of the indices appears more than once. Such a form I will call an *$R(s)$-form*. Thus, the index set corresponding to the s matrices in the $R(s)$-form is a proper subset of $\{1, 2, \ldots, s\}$. Examples of $A(3)$- and $R(3)$-forms are:
$A(3)$: $(\mathrm{tr}\, A_1)\,\mathrm{tr}\,(A_2 A_3)$, $(\mathrm{tr}\, A_1)(\mathrm{tr}\, A_2)(\mathrm{tr}\, A_3)$, $\mathrm{tr}\,(A_1 A_2 A_3)$;
$R(3)$: $(\mathrm{tr}\, A_3)\,\mathrm{tr}\,(A_2 A_3)$, $(\mathrm{tr}\, A_2)(\mathrm{tr}\, A_1)^2$, $\mathrm{tr}\, A_2^3$.
Now, by symmetry, the sum of all the $A(s)$-forms in $\mathrm{E} \prod_{j=1}^{s} u'A_j u$ is an $A(s)$-polynomial. Hence, I may write

$$\mathrm{E} \prod_{j=1}^{s} u'A_j u = A(s) + R(s),$$

where $A(s)$ is the $A(s)$-polynomial and $R(s)$ contains the $R(s)$-forms. Further, $\mathrm{E} \prod_{j=1}^{s} u'A_j u$ is homogeneous in the A_j $(j = 1 \ldots s)$, since, for all λ,

$$\mathrm{E} \left[u'(\lambda A_i) u \prod_{\substack{j=1 \\ j \neq i}}^{s} (u'A_j u) \right] = \lambda \mathrm{E} \prod_{j=1}^{s} u'A_j u.$$

Similarly, $A(s)$ is homogeneous in the A_j. Hence, $R(s)$ is homogeneous in the A_j. But this can only be if $R(s) \equiv 0$. This concludes the first part of the theorem.

To prove the second part, write

$$\mathrm{E} \prod_{j=1}^{s} \varepsilon' A_j \varepsilon = \mathrm{E} \prod_{j=1}^{s} u' V^{\frac{1}{2}} A_j V^{\frac{1}{2}} u = \mathrm{E} \prod_{j=1}^{s} u' \bar{A}_j u,$$

where $\bar{A}_j = V^{\frac{1}{2}} A_j V^{\frac{1}{2}}$. Then, subsituting $\bar{A}_j$ in the $A(s)$-polynomial and observing that $\operatorname{tr}(V^{\frac{1}{2}} A_j V^{\frac{1}{2}} V^{\frac{1}{2}} A_h V^{\frac{1}{2}}) = \operatorname{tr}(A_j V A_h V)$ concludes the proof.

REMARK

The required expectation is found as follows: First write down the $A(s)$-polynomial, then put $A_1 = A_2 = \ldots = A_s$, and use the known coefficients of $\mathrm{E}(u'Au)^s$ [see lemma 2.3] to determine the coefficients of the $A(s)$-polynomial. For example, the coefficients in the $A(2)$-, $A(3)$-, and $A(4)$-polynomials given under definition 3.4 are

$A(2)$: $\nu_1 = 1$, $\nu_2 = 2$,

$A(3)$: $\nu_1 = 1$, $\nu_2 = 2$, $\nu_3 = 8$,

$A(4)$: $\nu_1 = 1$, $\nu_2 = 8$, $\nu_3 = 4$, $\nu_4 = 2$, $\nu_5 = 16$.

6 Some applications: skewness and kurtosis

Let me now give a few examples. Let x be a stochastic variable, $\mathrm{E}x = \mu$, $\mathrm{E}(x-\mu)^2 = \sigma^2$. Clearly, μ and σ^2 are measures of location and dispersion. The ratio of the third moment around the mean to the third power of the standard deviation,

$$\beta = \sigma^{-3}\mathrm{E}(x-\mu)^3 \tag{6.1}$$

is often used as a measure for the degree of nonsymmetry (*skewness*) of a distribution. For distributions with a long tail in positive direction it takes positive values; for those with a long tail in the opposite direction it takes a negative value. If the distribution is symmetrical, β vanishes. The opposite, however, is not true, so the value of β must be interpreted with some caution. Another statistic is

$$\gamma = \sigma^{-4}\mathrm{E}(x-\mu)^4 - 3, \tag{6.2}$$

which is used as a measure of the degree of flattening of a frequency curve near its centre (*kurtosis*). Its value is usually positive when the distribution has long thick tails and sharp peaks relative to the normal distribution. Note that for the normal distribution, $\beta = \gamma = 0$.

Lemma 6.1

Let A be a symmetric matrix, and $\varepsilon \sim N_n(0, V)$, V positive definite. Define $y = \varepsilon' A \varepsilon$. The expectation, variance, skewness, and kurtosis of y are:

$$\begin{aligned}
\mu(y) &= \operatorname{tr} AV \\
\sigma^2(y) &= 2\operatorname{tr}(AV)^2 \\
\beta(y) &= 2\sqrt{2}\operatorname{tr}(AV)^3\{\operatorname{tr}(AV)^2\}^{-3/2} \\
\gamma(y) &= 12\operatorname{tr}(AV)^4\{\operatorname{tr}(AV)^2\}^{-2}.
\end{aligned}$$

Proof

It follows from lemma 2.3 that $\mu = \mathrm{E}y = \mathrm{tr}\, AV$, $\mathrm{E}(y-\mu)^2 = 2\mathrm{tr}\,(AV)^2$, $\mathrm{E}(y-\mu)^3 = 8\mathrm{tr}\,(AV)^3$, and $\mathrm{E}(y-\mu)^4 = 48\mathrm{tr}\,(AV)^4 + 12(\mathrm{tr}\,(AV)^2)^2$. Hence, by definitions (6.1) and (6.2) the result follows.

Notice that the kurtosis $\gamma(y)$ is always positive which indicates that the frequency curve of $\varepsilon' A\varepsilon$ is likely to be more tall and slim than the normal curve in the neighborhood of the mode.

Similar results can be found in NAGAR's case, i.e. for the product of two quadratic forms. The formulae, however, are more complicated.

Lemma 6.2

Let A and B be symmetric matrices, and $\varepsilon \sim N_n(0, V)$, V positive definite. Define $z = \varepsilon' A\varepsilon \cdot \varepsilon' B\varepsilon$. The expectation and variance of z are:

$$\mu(z) = (\mathrm{tr}\, AV)(\mathrm{tr}\, BV) + 2\mathrm{tr}\,(AVBV),$$

$$\begin{aligned}\sigma^2(z) = {} & 32\mathrm{tr}\,[(AV)^2(BV)^2] + 16[\mathrm{tr}\,(AVBV)^2 + (\mathrm{tr}\, AV)(\mathrm{tr}\, AV(BV)^2) \\ & + (\mathrm{tr}\, BV)(\mathrm{tr}\,(AV)^2 BV)] + 4[\mathrm{tr}\,(AV)^2\, \mathrm{tr}\,(BV)^2 + (\mathrm{tr}\, AVBV)^2 \\ & + (\mathrm{tr}\, AV)(\mathrm{tr}\, BV)(\mathrm{tr}\, AVBV)] + 2[(\mathrm{tr}\, AV)^2\, \mathrm{tr}\,(BV)^2 + (\mathrm{tr}\, BV)^2\, \mathrm{tr}\,(AV)^2].\end{aligned}$$

Proof

Let $u \sim N_n(0, I)$. Applying the formulae under definition 3.4 (with the known coefficients given in the remark under Theorem 5.1), it is easily seen that

$$\mathrm{E}u'Au \cdot u'Bu = (\mathrm{tr}\, A)(\mathrm{tr}\, B) + 2\,\mathrm{tr}\,(AB)$$

and

$$\begin{aligned}\mathrm{E}(u'Au \cdot u'Bu)^2 = {} & (\mathrm{tr}\, A)^2(\mathrm{tr}\, B)^2 + 16[(\mathrm{tr}\, A)(\mathrm{tr}\, AB^2) + (\mathrm{tr}\, B)(\mathrm{tr}\, A^2 B)] \\ & + 4[(\mathrm{tr}\ A^2)\ (\mathrm{tr}\ B^2) + 2(\mathrm{tr}\ AB)^2] + 2[(\mathrm{tr}\ A)^2(\mathrm{tr}\ B^2) \\ & + 4(\mathrm{tr}\, A)(\mathrm{tr}\, B)(\mathrm{tr}\, AB) + (\mathrm{tr}\, B)^2(\mathrm{tr}\, A^2)] + 16[\mathrm{tr}\,(AB)^2 \\ & + 2(\mathrm{tr}\, A^2B^2)].\end{aligned}$$

Hence, the variance of $u'Au \cdot u'Bu$ is

$$\begin{aligned}\sigma^2(u'Au \cdot u'Bu) = {} & \mathrm{E}(u'Au \cdot u'Bu)^2 - [\mathrm{E}(u'Au \cdot u'Bu)]^2 = 32\,\mathrm{tr}\,(A^2B^2) \\ & + 16[\mathrm{tr}\,(AB)^2 + (\mathrm{tr}\, A)(\mathrm{tr}\, AB^2) + (\mathrm{tr}\, B)(\mathrm{tr}\, A^2B)] \\ & + 4[(\mathrm{tr}\, A^2)(\mathrm{tr}\, B^2) + (\mathrm{tr}\, AB)^2 + (\mathrm{tr}\, A)(\mathrm{tr}\, B)(\mathrm{tr}\, AB)] \\ & + 2[(\mathrm{tr}\, A)^2(\mathrm{tr}\, B^2) + (\mathrm{tr}\, B)^2(\mathrm{tr}\, A^2)].\end{aligned}$$

Substituting AV for A, and BV for B gives the desired results.

Acknowledgements

I am grateful to PETER VAN EMDE BOAS, RISTO D. H. HEIJMANS, ROALD RAMER, two referees and the editor of Statistica for their helpful comments.

References

[1] CRAMÉR, H., Mathematical methods of statistics. Princeton University Press, 1946.
[2] KENDALL, M. G. and A. STUART, The advanced theory of statistics, Vol. I, 3rd edition. Charles Griffin & Co., London, 1947.
[3] MAGNUS, J. R. and H. NEUDECKER, The commutation matrix: some properties and applications. Annals of Statistics (1979), 7, forthcoming.
[4] NAGAR, A. L., The bias and moment matrix of the general k-class estimators of the parameters in simultaneous equations. Econometrica (1959), **27**, 575–595.
[5] NEUDECKER, H., The Kronecker matrix product and some of its applications in econometrics. Statistica Neerlandica (1968), **22**, 69–82.
[6] OGASAWARA, T. and M. TAKAHASHI, Independence of quadratic quantities in a normal system. J. Sci. Hiroshima University (1951), **15**, 1–9.
[7] RAO, C. R. and S. K. MITRA, Generalized inverse of matrices and applications. Wiley, New York, 1971.