Metrika, Volume 33, 1986, page 247–251

A Note on Some Wishart Expectations[1]

By R. J. Muirhead[2]

Summary: In a recent paper Sharma and Krishnamoorthy (1984) used a complicated decision-theoretic argument to derive an identity involving expectations taken with respect to the Wishart distribution $W_m(n, I)$. A more general result, proved using an elementary moment generating function argument, and some applications, are given in this paper.

1 An Expectation Identity

Let the random $m \times m$ matrix S have the Wishart distribution $W_m(n, \Sigma)$, with probability density function

$$c_{m,n}(\det \ \Sigma)^{-n/2} \exp(-{}^1\!/_2 \operatorname{tr} \Sigma^{-1} S)(\det S)^{1/2(n-m-1)}, \quad S > 0, \ \Sigma > 0, \ 0, \ n > m-1$$

where

$$c_{m,n}^{-1} = 2^{mn/2} \Gamma_m({}^1\!/_2 \, n),$$

with

$$\Gamma_m(a) = \pi^{1/4m(m-1)} \prod_{i=1}^{m} \Gamma(a - {}^1\!/_2(i-1)).$$

Using an innovative, but rather complicated and involved decision-theoretic argument, Sharma and Krishnamoorthy (1984) proved that when $\Sigma = I_m$,

$$E[(\operatorname{tr} S)^2 \operatorname{tr}(S^\alpha)] = (mn + 2 + 2\alpha) E[\operatorname{tr} S \operatorname{tr}(S^\alpha)], \tag{1}$$

an identity which holds for all α for which the expectations exist. Here we give an elementary proof of a more general result which yields (1) as a special case. The general result is given in the following theorem.

[1] This work was supported by the National Science Foundation.

[2] Prof. Robb J. Muirhead, Department of Statistics, University of Michigan, 419 South State Street, Ann Arbor, Michigan 48109-1027, USA.

Theorem: Suppose that $S \sim W_m(n, \Sigma)$. Let $h(S)$ be a real-valued measurable function of S such that the function $f(t;S) = h(t^{-1}S)$, $t > 0$, is differentiable at $t = 1$. Let $f'(t;S) = \frac{\partial}{\partial t} f(t;S)$. Then

$$E[\text{tr}\,(\Sigma^{-1}S)h(S)] = mnE[h(S)] - 2E[f'(1;S)] \tag{2}$$

provided the expectations involved exist.

Proof: For $t > 0$ define the function $g(t)$ as

$$g(t) = c_{m,n}(\det\,\Sigma)^{-n/2} t^{mn/2} \int_{S>0} \exp\left(-\frac{t}{2}\,\text{tr}\,\Sigma^{-1}S\right)(\det S)^{1/2(n-m-1)}h(S)dS, \tag{3}$$

and note that $g(1) = E[h(S)]$.

Differentiating (3) with respect to t (justified by dominated convergence provided $E[\text{tr}\,(\Sigma^{-1}S)h(S)]$ exists) and putting $t = 1$ gives

$$g'(1) = {}^1\!/\!_2\, mnE[h(S)] - {}^1\!/\!_2\, E[\text{tr}\,(\Sigma^{-1}S)h(S)]. \tag{4}$$

Now put $X = tS$ in (3); then $g(t)$ can be written alternatively as

$$g(t) = c_{m,n}(\det\,\Sigma)^{-n/2} \int_{X>0} \exp\,(-{}^1\!/\!_2\,\text{tr}\,\Sigma^{-1}X)(\det X)^{1/2(n-m-1)}f(t;X)dX,$$

from which it follows that

$$g'(1) = E[f'(1;S)]. \tag{5}$$

Equating (4) and (5) gives the desired result (2) and completes the proof.

In many interesting applications the function $h(\cdot)$ has the property that, for $x > 0$, $h(xS) = x^l h(S)$ for some real l. Then $f(t;S) = h(t^{-1}S) = t^{-l}h(S)$, so that

$$f'(1;S) = -lh(S).$$

This yields the following result.

Corollary: If $h(xS) = x^l h(S)$ for some l then

$$E[\text{tr}\,(\Sigma^{-1}S)h(S)] = (mn + 2l)E[h(S)]. \tag{6}$$

The identity (1) of Sharma and Krishnamoorthy (1984) follows immediately from (6) by taking $\Sigma = I$ and $h(S) = \text{tr}\,S\,\text{tr}\,(S^\alpha)$, so that $l = \alpha + 1$. Another identity, used by

Efron and Morris (1976) in the context of decision-theoretic estimation of Σ^{-1}, is

$$E\left[\frac{\operatorname{tr}\Sigma^{-1}S}{\operatorname{tr}S}\right]=(mn-2)E\left[\frac{1}{\operatorname{tr}S}\right];$$

this is a special case of (6) with $h(S)=(\operatorname{tr}S)^{-1}$, so that $l=-1$.

2 Applications

Many interesting expectations can be evaluated using (6). Some of these, in which k and r are nonnegative integers and

$$(a)_k=a(a+1)\ldots(a+k-1),$$

are:

$$E[(\operatorname{tr}\Sigma^{-1}S)^k]=2^k(\tfrac{1}{2}mn)_k, \tag{7}$$

$$E[(\operatorname{tr}\Sigma^{-1}S)^{-k}]=(-\tfrac{1}{2})^k/(-\tfrac{1}{2}mn+1)_k, \quad (2k<mn) \tag{8}$$

$$E[(\operatorname{tr}\Sigma^{-1}S)^k\operatorname{tr}S]=2^k(\tfrac{1}{2}mn+1)_k n\operatorname{tr}\Sigma, \tag{9}$$

$$E[(\operatorname{tr}\Sigma^{-1}S)^k\operatorname{tr}S^{-1}]=2^k(\tfrac{1}{2}mn-1)_k\operatorname{tr}\Sigma^{-1}/(n-m-1), \quad (n>m+1) \tag{10}$$

$$E[(\operatorname{tr}\Sigma^{-1}S)^k\operatorname{tr}\Sigma S^{-1}]=2^k(\tfrac{1}{2}mn-1)_k m/(n-m-1), \quad (n>m+1) \tag{11}$$

$$E[(\operatorname{tr}\Sigma^{-1}S)^k(\det S)^r]=2^{mr+k}(\tfrac{1}{2}mn+rm)_k\frac{\Gamma_m(\tfrac{1}{2}n+r)}{\Gamma_m(\tfrac{1}{2}n)}\cdot(\det\Sigma)^r. \tag{12}$$

These may all be derived using essentially similar arguments and known elementary properties of Wishart matrices. For example, (7) is proved in the following way. Put $h(S)=\operatorname{tr}(\Sigma^{-1/2}S\Sigma^{-1/2})=\operatorname{tr}(\Sigma^{-1}S)$ in (6), so that $l=1$. This gives

$$E[(\operatorname{tr}\Sigma^{-1}S)^2]=(mn+2)E[\operatorname{tr}(\Sigma^{-1}S)]=(mn+2)mn$$

where we have used the fact that $E(S)=n\Sigma$. Next, taking $h(S)=(\operatorname{tr}\Sigma^{-1}S)^2$, with $l=2$, gives

$$E[(\operatorname{tr}\Sigma^{-1}S)^3]=(mn+4)E[(\operatorname{tr}\Sigma^{-1}S)^2]=(mn+4)(mn+2)mn.$$

The result (7) for arbitrary k follows trivially by induction. To prove (8) note that taking $h(S)=1/\operatorname{tr}\Sigma^{-1}S$ in (6) gives $(mn-2)E\left[\dfrac{1}{\operatorname{tr}\Sigma^{-1}S}\right]=1$.

Next, taking $h(S) = (\text{tr }\Sigma^{-1}S)^{-2}$ in (6) gives

$$(mn-4)E[(\text{tr }\Sigma^{-1}S)^{-2}] = E[(\text{tr }\Sigma^{-1}S)^{-1}] = 1/(mn-2),$$

and the rest of the argument is obvious. The proofs of the other identities are similar. Note that to derive (10), (11), and (12) we need the known results

$$E(S^{-1}) = \frac{1}{n-m-1}\Sigma^{-1},$$

$$E[(\det S)^r] = 2^{mr}\frac{\Gamma_m(\tfrac{1}{2}n+r)}{\Gamma_m(\tfrac{1}{2}n)}(\det \Sigma)^r.$$

We conclude by giving two expectations involving zonal polynomials. Let $C_\kappa(S)$ denote the zonal polynomial of S corresponding to the partition $\kappa = (k_1, k_2, \ldots, k_m)$ of the integer k ($k_1 \geqslant k_2 \geqslant \ldots \geqslant k_m \geqslant 0$) (see e.g. James 1964 or Muirhead 1982, Chapter 7), and let

$$(a)_\kappa = \prod_{i=1}^{m}(a - \tfrac{1}{2}(i-1))_{k_i}.$$

The following expectations, in which B is a non-random mxm symmetric matrix, hold:

$$E[(\text{tr }\Sigma^{-1}S)^r C_\kappa(SB)] = (\tfrac{1}{2}mn+k)_r 2^{k+r}(\tfrac{1}{2}n)_\kappa C_\kappa(B\Sigma), \tag{13}$$

$$E[(\text{tr }\Sigma^{-1}S)^{-r} C_\kappa(SB)] = \frac{(-1)^r 2^{k-r}}{(-\tfrac{1}{2}mn-k+1)_r}(\tfrac{1}{2}n)_\kappa C_\kappa(B\Sigma) \quad (r < \tfrac{1}{2}mn+k). \tag{14}$$

These may be derived from (6) in a similar way to (7)–(12) using the known result that

$$E[C_\kappa(SB)] = 2^k(\tfrac{1}{2}n)_\kappa C_\kappa(B\Sigma).$$

(see e.g. Muirhead 1982, p. 251).

References

Efron B, Morris C (1976) Multivariate empirical Bayes and estimation of covariance matrices. Ann Statist 4:22–32

James AT (1964) Distributions of matrix variates and latent roots derived from normal samples. Ann Math Statist 35:475–501

Muirhead RJ (1982) Aspects of multivariate statistical theory. John Wiley & Sons, New York

Sharma D, Krishnamoorthy K (1984) An identity involving a Wishart matrix. To appear in Metrika

Received July 10, 1984
(Revised version August 6, 1984)