An Identity Concerning a Wishart Matrix.

Sharma, D.; Krishnamoorthy, K.

pp. 65 - 68

Metrika, Volume 33, 1986, page 65–68

An Identity Concerning a Wishart Matrix

By D. Sharma and K. Krishnamoorthy[1]

Summary: Let S be a $p \times p$ Wishart matrix with parameters n and Σ. For a rational number $\alpha = r/s$ with r and s integers and s positive, let S^α denote a positive definite matrix such that $(S^\alpha)^s = S^r$. Using a decision theoretic argument, we prove that $E[(\operatorname{tr} S)^2 \operatorname{tr} S^\alpha] = (np + 2 + 2\alpha) E[\operatorname{tr} S \operatorname{tr} S^\alpha]$ when $\Sigma = I$ and $np + 2 + 2\alpha$ is positive.

Let the $p \times p$ random matrix S follow a Wishart distribution with parameters n and Σ, with its density proportional to

$$|S|^{1/2(n-p-1)} e^{-1/2 \operatorname{tr} \Sigma^{-1} S}, \quad S > 0, \Sigma > 0, n \geqslant p.$$

Then we prove the following identity:

For $\Sigma = I$, $np + 2 + 2\alpha > 0$, $\alpha = r/s$ with r and s integers and s positive, and $S^\alpha > 0$ such that $(S^\alpha)^s = S^r$,

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

For the proof, we take a loss function $L(\Sigma, \hat{\Sigma}) = \operatorname{tr}(\hat{\Sigma} - \Sigma)^2 S^\alpha / \operatorname{tr} \Sigma^{2+\alpha}$, a class of estimators $D = \{b(\operatorname{tr} S)I,\ b$ a positive number$\}$, and a subset of the parameter space $\Omega_0 = \{\Sigma : \Sigma^{-1} = KI,\ K$ a positive number$\}$, and notice that while the members of D are scale equivariant the loss function is scale invariant. As Ω_0 is an orbit of the parameter space under the scale group of transformations, the risk of $b(\operatorname{tr} S)I$ is constant for Σ in Ω_0. Let b_0 be such that the risk of $b(\operatorname{tr} S)I$ is minimized for $b = b_0$ when $\Sigma \in \Omega_0$. If we can find an estimator $b_1(\operatorname{tr} S)I$, which is minimax for $\Sigma \in \Omega_0$, then b_0 and b_1 should be equal and this is what gives us the identity (1).

Now the risk of $b(\operatorname{tr} S)I$ when $\Sigma \in \Omega_0$ is

$$p^{-1}[b^2 E_{\Sigma=I}(\operatorname{tr} S)^2 \operatorname{tr} S^\alpha + E_{\Sigma=I} \operatorname{tr} S^\alpha - 2b E_{\Sigma=I}(\operatorname{tr} S) \operatorname{tr} S^\alpha]$$

so that the minimizing b_0 is

$$E_{\Sigma=I}(\operatorname{tr} S) \operatorname{tr} S^\alpha / E_{\Sigma=I}(\operatorname{tr} S)^2 \operatorname{tr} S^\alpha \tag{2}$$

---

[1] D. Sharma and K. Krishnamoorthy, Department of Mathematics, Indian Institute of Technology, Kanpur.

To obtain a minimax estimator of the form $b(\text{tr}\, S)I$ for $\Sigma \in \Omega_0$, we shall find a sequence of Bayes estimators with the corresponding sequence of Bayes risks tending to the risk of an estimator $b_1(\text{tr}\, S)I$. This convergence then implies the minimaxity of $b_1(\text{tr}\, S)I$ [see, for example, Ferguson, p. 90]. The following lemma is needed for this purpose.

Lemma: For a given prior distribution of Σ, let $E_{\Sigma|S}$ denote the expectation with respect to the posterior distribution of Σ. Then, for the loss function $L(\Sigma, \hat{\Sigma})$, the Bayes estimator of Σ is

$$\tilde{\Sigma} = [E_{\Sigma|S}(\text{tr}\, \Sigma^{2+\alpha})^{-1}]^{-1} E_{\Sigma|S}[\Sigma(\text{tr}\, \Sigma^{2+\alpha})^{-1}].$$

Proof: The posterior risk of $\hat{\Sigma}$ is

$$\begin{aligned}
&E_{\Sigma|S}\, \text{tr}(\hat{\Sigma} - \Sigma)^2 S^\alpha/\text{tr}\, \Sigma^{2+\alpha} \\
&= E_{\Sigma|S}\, \text{tr}(\hat{\Sigma}^2 + \Sigma^2 - \hat{\Sigma}\Sigma - \Sigma\hat{\Sigma})S^\alpha/\text{tr}\, \Sigma^{2+\alpha} \\
&= \text{tr}\, \hat{\Sigma}^2 S^\alpha E_{\Sigma|S}(\text{tr}\, \Sigma^{2+\alpha})^{-1} + E_{\Sigma|S}(\text{tr}\, \Sigma^2 S^\alpha/\text{tr}\, \Sigma^{2+\alpha}) \\
&- \text{tr}\, \hat{\Sigma} E_{\Sigma|S}(\Sigma/\text{tr}\, \Sigma^{2+\alpha})S^\alpha - \text{tr}\, E_{\Sigma|S}(\Sigma/\text{tr}\, \Sigma^{2+\alpha})\hat{\Sigma}S^\alpha \\
&= [E_{\Sigma|S}(\text{tr}\, \Sigma^{2+\alpha})^{-1}]\, [\text{tr}(\hat{\Sigma} - \tilde{\Sigma})^2 S^\alpha - \text{tr}\, \tilde{\Sigma}^2 S^\alpha] \\
&+ E_{\Sigma|S}(\text{tr}\, \Sigma^2 S^\alpha/\text{tr}\, \Sigma^{2+\alpha})
\end{aligned} \tag{3}$$

As $\hat{\Sigma} = \tilde{\Sigma}$ minimizes (3), $\tilde{\Sigma}$ is the Bayes estimator of Σ.

Incidentally, the lemma also follows from Eaton's proposition 2.1 [p. 2].

Next, we prove the following minimaxity result.

Theorem: If $\alpha, L(\Sigma, \hat{\Sigma})$ and Ω_0 are as defined earlier and $np + 2 + 2\alpha > 0$, then the estimator $(\text{tr}\, S)I/(np + 2 + 2\alpha)$ is minimax for $\Sigma \in \Omega_0$.

Proof: Let K be as in the definition of Ω_0. Following Efron/Morris [p. 119], take the prior density of K to be $k^{-a}(1-a)$ for $0 < k \leqslant 1, a < 1$. This determines a prior distribution of $\Sigma \in \Omega_0$. From the above lemma, the Bayes estimator of Σ is

$$\hat{\Sigma}_a = [\int_0^1 e^{-1/2k\, \text{tr}\, S} k^{1/2np - a + \alpha + 1}\, dk / \int_0^1 e^{-1/2k\, \text{tr}\, S} k^{1/2np - a + \alpha + 2}\, dk] I$$

$$= \frac{\text{tr}\, S}{np - 2a + 2\alpha + 4} [G_{1/2np - a + \alpha + 2}(1/2\ \text{tr}\, S)/G_{1/2np - a + \alpha + 3}(1/2\ \text{tr}\, S)] I,$$

where

$$G_\beta(x) = \int_0^x e^{-t} t^{\beta - 1}\, dt/\Gamma(\beta).$$

Let $\tau(S) = \dfrac{np + 2 + 2\alpha}{np - 2a + 2\alpha + 4} \cdot \dfrac{G_{1/2np - a + \alpha + 2}(1/2 \operatorname{tr} S)}{G_{1/2np - a + \alpha + 3}(1/2 \operatorname{tr} S)}$ and $\hat{\Sigma}_1 = (\operatorname{tr} S)I/(np + 2 + 2\alpha)$.

Then, $\hat{\Sigma}_a = \tau(S)\hat{\Sigma}_1$ and, as pointed out in Efron/Morris (Ibid.), $\tau(S)$ increases monotonically from 0 to $(np + 2 + 2\alpha)/(np - 2a + 2\alpha + 4)$ as $\operatorname{tr} S$ increases from 0 to ∞. The Bayes risk of $\hat{\Sigma}_a$ is the expectation of

$$\operatorname{tr}\left[\left\{\tau^2(S)\,\hat{\Sigma}_1^2 + \frac{I}{K^2} - \frac{2\tau(S)\hat{\Sigma}_1}{K}\right\} S^{\alpha}\right] \frac{K^{2+\alpha}}{p}$$

under the joint distribution of S and K. Now, if a tends to $1, K$ converges in distribution to 0, so that $\operatorname{tr} S$ converges in distribution to ∞. As a consequence, $\tau(S)$ converges in distribution to 1 and the Bayes risk of $\hat{\Sigma}_a$ tends to the risk of $\hat{\Sigma}_1$ at $\Sigma \in \Omega_0$. This proves the theorem.

The identity (1) now follows from (2) and the theorem. Some consequences of (1) are, for $\Sigma = I$,

(i) $E(\operatorname{tr} S)^2 = (np + 2)E(\operatorname{tr} S) = np(np + 2)$

(ii) $E(\operatorname{tr} S)^3 = (np + 4)E(\operatorname{tr} S)^2 = np(np + 2)(np + 4)$

(iii) $E(\operatorname{tr} S)^2(\operatorname{tr} S^{-1}) = np\,E(\operatorname{tr} S)(\operatorname{tr} S^{-1})$

$\qquad = E(\operatorname{tr} S)E(\operatorname{tr} S)(\operatorname{tr} S^{-1})$

(iv) $E(\operatorname{tr} S)^2(\operatorname{tr} S^{-2}) = (np - 2)E(\operatorname{tr} S)(\operatorname{tr} S^{-2})$

(v) $E(\operatorname{tr} S)^2(\operatorname{tr} S^{1/2}) = (np + 3)E(\operatorname{tr} S)(\operatorname{tr} S^{1/2})$.

Whereas (i) and (ii) are trivial, (iii)[2] and (iv) can also be obtained in a straight-forward manner using the relations

$$ES^{-1}e^{-\theta \operatorname{tr} S} = |I + 2\theta\Sigma|^{-1/2n}(\Sigma^{-1} + 2\theta I)/(n - p - 1)$$

and

$$ES^{-2}e^{-\theta \operatorname{tr} S} = |I + 2\theta\,\Sigma|^{-1/2n}\left[\frac{(\Sigma^{-1} + 2\theta I)^2}{(n - p)(n - p - 3)} + \frac{[\operatorname{tr}(\Sigma^{-1} + 2\theta I)](\Sigma^{-1} + 2\theta I)}{(n - p)(n - p - 1)(n - p - 3)}\right]$$

(Haff).

However, verification of (1) for large integral values of α and for fractional values of α, for example, the relation (v), may not be that simple. The relation (1) is quite

---

[2] The alternative proof of (iii) communicated to the author by Professor C. G. Khatri.

general; it is valid for any α for which S^{α} is well defined and is positive definite (see the proof of the lemma).

It is also our feeling that the type of argument used here can lead, through a proper choice of loss functions, to other interesting identities.

References

Eaton ML (1970) Some problems in covariance estimation. Standford University Technical Report, No 49

Efron B, Morris C (1973) Stein's estimation rule and its competitors – an empirical Bayes approach. J Amer Statist Assoc 68: 117–130

Ferguson TS (1967) Mathematical statistics: a decision theoretic approach. Academic, New York

Haff LR (1980) Identities for the inverse wishart distribution and their relation to the Wishart case. University of California, San Diego Technical Report

Received November 16, 1982
(Revised version May 16, 1983)