ON CERTAIN CHARACTERISTICS OF THE DISTRIBUTION OF THE LATENT ROOTS OF A SYMMETRIC RANDOM MATRIX UNDER GENERAL CONDITIONS[1]

By H. R. van der Vaart

Leiden University, Netherlands

1. Summary. Under certain conditions, to be specified in Theorems 2 and 4, the latent roots of the symmetric random matrix F with $\mathcal{E}F = \Phi$ are biased estimators of the latent roots of Φ; the smallest (largest) root is negatively (positively) biased. Here bias includes both expectation-bias and median-bias. Further properties of the distribution of the latent roots are given, among them some relations between covariances of the latent roots, covariances of elements of F, and the amounts of expectation-bias of the latent roots. Also, a sufficient condition is given for a certain type of symmetry in the joint distribution of the latent roots. For applications of the theory presented in this paper to the theory of response surface estimation see van der Vaart [9].

2. Introduction, notations, definitions, statement of the problem. We shall use Latin letters for random variables, Greek letters for parameters, small letters for real-valued variables, capital letters for square matrices (examples: f_{ij}, l_h, u_{gh}, v_{kh} real-valued random variables; F, L, U, V random matrices; φ_{ij}, λ_h, v_{gh}, γ_h real-valued parameters; Φ, Λ, Υ, Γ matrices consisting of parameters). Let u_{gh} be an element of a matrix U, then $u_{\cdot h}$ will be used for the hth column vector in U so that $u_{\cdot h}$ is a $k \times 1$ matrix whose transpose, $u'_{\cdot h}$, is the hth row of U'. Finally, a symbol like $\mathcal{E}(F)$, the expectation of a matrix, will denote the matrix of the expectations,

$$(2.1) \qquad \mathcal{E}(F) = \|\mathcal{E}f_{ij}\| \begin{matrix} i = 1 \cdots k \\ j = 1 \cdots k \end{matrix},$$

where the superscript $i = 1 \cdots k$ denotes row numbers, the subscript $j = 1 \cdots k$ denotes column numbers (a good example of the usefulness of this compact notation is the defining equation (4.21) in Section 4).

Now consider a $k \times k$ matrix M with real elements m_{ij}. If a probability measure is defined in $\mathfrak{M}$, the set of possible k^2-tuples $(m_{11}, \cdots, m_{kk})$, then the matrix M will be called a random matrix. This probability measure is singular (relative to k^2-dimensional Lebesgue measure in $\mathfrak{M}$), if a subset of $\mathfrak{M}$ exists with Lebesgue measure zero and probability measure one (cf., p. 30 of Saks [6], or p. 611 of Doob [2]). This subset, defined up to a set of k^2-dimensional Lebesgue measure zero, will be denoted by the term M-space, and the probability distri-

Received February 4, 1960; revised March 15, 1961.

[1] This paper reports on research carried out while the author was visiting at the Institute of Statistics, Raleigh. This research was supported by the Office of Ordnance Research, U.S. Army, under contract DA-36-034-ORD-1517(RD).

bution over M-space will be denoted by the term (*probability*) *distribution of* M.

Examples of singular probability measures in $\mathfrak{M}$: M may be prescribed to be orthogonal (M-space is then $\frac{1}{2}k(k - 1)$-dimensional) or symmetric (M-space is $\frac{1}{2}k(k + 1)$-dimensional). In the latter case, M-space could be represented as a subset of $\mathfrak{M}$ by the equations

$$(2.2) \qquad m_{ij} = t_{ij} \quad (j \geqq i; i = 1, \cdots, k); \qquad m_{ij} = t_{ji} \quad (j < i; i = 1, \cdots, k).$$

We shall, however, define M-space to be the projection of this subset on $(m_{11}, \cdots, m_{1k}, m_{22}, \cdots, m_{2k}, \cdots, m_{k-1\,k-1}, m_{k-1\,k}, m_{kk})$-space. We shall call the *distribution* of a $k \times k$ *symmetric* random *matrix continuous* if it is absolutely continuous relative to $\frac{1}{2}k(k + 1)$-dimensional Lebesgue measure on $m_{ij}(1 \leqq i \leqq j \leqq k)$-space.

Two random matrices M' and M'' are called (stochastically) *equivalent*,

$$(2.3a) \qquad M' \sim M'',$$

if $M' = M''$ with probability one (cf., p. 33 of Kolmogorov [4]). Let $f(M)$ be a real-valued function of the matrix M; if M' and M'' are equivalent, then

$$(2.3b) \qquad \mathcal{E}f(M') = \mathcal{E}f(M''),$$

provided these expectations exist.

Our problem now is this. Let Φ be a real symmetric $k \times k$ matrix. Let F be a real symmetric $k \times k$ random matrix, which is continuously distributed and satisfies

$$(2.4) \qquad \mathcal{E}(F) = \Phi.$$

Then the k latent roots, l_h of F and λ_h of Φ, are real $(h = 1, \cdots, k)$. *Assign the subscripts* in such a way that

$$(2.5) \qquad l_1 \leqq l_2 \leqq \cdots \leqq l_k\,; \lambda_1 \leqq \lambda_2 \leqq \cdots \leqq \lambda_k\,.$$

Define diagonal matrices L and Λ by

$$(2.6) \qquad L = \|l_g\delta_{gh}\| \begin{matrix} g = 1 \cdots k \\ h = 1 \cdots k \end{matrix}; \qquad \Lambda = \|\lambda_g\delta_{gh}\| \begin{matrix} g = 1 \cdots k \\ h = 1 \cdots k \end{matrix}.$$

Our aim is to investigate the *distribution of* L. Note that, although two or more roots λ_h may be exactly equal, the probability that two or more roots l_h be exactly equal, is always zero, since F is continuously distributed.

3. Definition of a few important matrices. As is well known, one can always construct a random orthogonal matrix U with real elements such that $FU = UL$. Here the column $u_{\cdot h}$ of U is the latent vector (= eigenvector) corresponding to the latent root l_h. If two or more latent roots are equal, $l_{h_1} = l_{h_2} = \cdots = l_{h_d}$, say, then the columns $u_{\cdot h_1}, u_{\cdot h_2}, \cdots, u_{\cdot h_d}$ of any orthogonal U with $FU = UL$ are a basis for the eigenspace corresponding to the latent root $l_{h_1} = \cdots = l_{h_d}$.

What we have said just now, may be repeated with Υ, Φ, Λ, v, λ instead of U, F, L, u, l. We list the following for reference:

$$\text{(3.1a)} \qquad \Phi = \Upsilon\Lambda\Upsilon', \qquad F = ULU',$$

$$\text{(3.1b)} \qquad \Phi\Upsilon = \Upsilon\Lambda, \qquad FU = UL,$$

$$\text{(3.1c)} \qquad \Upsilon'\Phi\Upsilon = \Lambda, \qquad U'FU = L,$$

$$\text{(3.1d)} \qquad \Upsilon'\Phi = \Lambda\Upsilon', \qquad U'F = LU'.$$

As the probability that two or more latent roots of F be equal, is zero, the columns $u_{\cdot h}$ of U are uniquely defined with probability one, except for their sign. As to this question of signs, for the purposes of the present paper it is sufficient to observe that it is possible to restrict this ambiguity in such a way that Det $(U) = +1$ throughout, and that, with L fixed and U variable, the matrix $F = ULU'$ runs just once through all values F that have the same matrix L of latent roots $l_1 \leqq \cdots \leqq l_k$. If two or more latent roots of Φ are equal, then in order to represent Φ by $\Upsilon\Lambda\Upsilon'$ it will suffice to choose just one orthogonal basis in the corresponding eigenspace.

U and Υ being defined, we will now define V by one of the equivalent relations

$$\text{(3.2)} \qquad \begin{matrix} V = \Upsilon'U, & \Upsilon V = U, & \Upsilon = UV', \\ V' = U'\Upsilon, & V'\Upsilon' = U', & \Upsilon' = VU'. \end{matrix}$$

Evidently V is an orthogonal random matrix with Det $(V) = +1$.

Finally, define

$$\text{(3.3)} \qquad \hat{L} = VLV'.$$

Equations (3.1c), (2.4), (2.1), (3.2), (3.1c) show that

$$\text{(3.4)} \quad \Lambda = \Upsilon'\Phi\Upsilon = \Upsilon'\cdot\mathcal{E}(F)\cdot\Upsilon = \mathcal{E}(\Upsilon'F\Upsilon) = \mathcal{E}(VU'FUV') = \mathcal{E}(VLV') = \mathcal{E}(\hat{L}).$$

Hence $\hat{L}$ would be an unbiased estimator of Λ if it were a statistic. Unfortunately, though U is a statistic, Υ is not. As Υ is usually unknown, V is not a statistic, nor is $\hat{L}$. Note that, whereas the off-diagonal elements of L equal zero by definition, the off-diagonal elements of $\hat{L}$ do not. Therefore one ought to write $\hat{l}_{gg}$ for the diagonal elements of $\hat{L}$. Yet, because of the analogy between $\hat{L}$ and L, it will sometimes be convenient to write $\hat{l}_g$ for $\hat{l}_{gg} (g = 1, \cdots, k)$. Furthermore note that $\hat{l}_1$ may well be larger than $\hat{l}_k$, whereas by definition $l_1 \leqq l_k$.

4. On certain characteristics of the distribution of L and $\hat{L}$. Equation (3.4) suggests that L may have certain undesirable features as an estimator of Λ. We are going to investigate the distribution of L (and $\hat{L}$). We will adhere throughout to the following assumption.

ASSUMPTION A. *The symmetric random matrix F is continuously distributed,*[2] *and* $\mathcal{E}(F) = \Phi$.

[2] The phrase "continuous distribution of a symmetric random matrix" has been defined in Section 2 as meaning *absolute* continuity with respect to a natural measure.

It is easy to prove the following theorem.

THEOREM 1.

$$(4.1)\qquad \mathcal{E}\,(\mathrm{tr}\,\hat{L}) = \mathcal{E}\,(\mathrm{tr}\,L) = \mathrm{tr}\,\Lambda = \mathrm{tr}\,\Phi = \mathcal{E}\,(\mathrm{tr}\,F).$$

PROOF. These equalities follow immediately from (3.4) and the following properties of traces:

$$(4.2)\qquad \mathrm{tr}\,\mathcal{E}(F) = \mathcal{E}\,(\mathrm{tr}\,F), \qquad \mathrm{tr}\,(AB) = \mathrm{tr}\,(BA).$$

For instance, $\mathrm{tr}\,(\hat{L}) = \mathrm{tr}\,(VLV') = \mathrm{tr}\,(V'VL) = \mathrm{tr}\,(L)$; hence $\mathcal{E}\,\mathrm{tr}\,(L) = \mathcal{E}\,\mathrm{tr}\,(\hat{L}) = \mathrm{tr}\,\mathcal{E}(\hat{L}) = \mathrm{tr}\,\Lambda$, etc. Proof completed.

Next we want to prove that L is a biased estimator of Λ in various respects. The following elegant argument was suggested to me by T. W. Anderson in a personal communication (February 1958). It is based on the following well known extremal property of the latent roots (see, for example, Satz 10, p. 292 of Gantmacher, [3]):

$$(4.3)\qquad l_k = \max\,(a'Fa) \geqq v'_{\cdot k}Fv_{\cdot k}\,,$$

where the maximum is taken over all vectors ($= k \times 1$ matrices) a with $a'a = 1$: hence the inequality (4.3) holds for all vectors $v_{\cdot k}$ in the eigenspace corresponding to the latent root λ_k of Φ. Inequality (4.3) yields

$$(4.4)\qquad \mathcal{E}l_k \geqq v'_{\cdot k}\cdot\mathcal{E}(F)\cdot v_{\cdot k} = v'_{\cdot k}\Phi v_{\cdot k} = \lambda_k\,.$$

A similar argument yields that $\mathcal{E}l_1 \leqq \lambda_1$.

Another consequence of (4.3) is that Med $l_k \geqq$ Med $(v'_{\cdot k}Fv_{\cdot k})$, and hence that

$$(4.5)\qquad \mathrm{Med}\,l_k \geqq \lambda_k\,, \quad \text{if} \quad \mathrm{Med}\,(v'_{\cdot k}Fv_{\cdot k}) \geqq \lambda_k\,.$$

A similar argument yields that Med $l_1 \leqq \lambda_1$ if Med $(v'_{\cdot 1}Fv_{\cdot 1}) \leqq \lambda_1$.

In order, however, to investigate conditions under which the inequality (4.5) is strict, and as a preliminary to a more detailed study of the distribution of L, we shall indicate a different method of proof. First we shall prove a lemma.

LEMMA 1. *Under the general assumption* A

$$(4.6)\qquad P(\hat{l}_1 - l_1 > 0) = 1, \qquad P(l_k - \hat{l}_k > 0) = 1.$$

PROOF. As V is orthogonal, we have that $\sum_{h=1}^{k} v_{1h}^2 = 1$, $\sum_{h=1}^{k} v_{kh}^2 = 1$. Hence, because of (3.3) we have

$$(4.7)\qquad \hat{l}_1 - l_1 = \sum_{h=1}^{k} v_{1h}^2(l_h - l_1), \qquad l_k - \hat{l}_k = \sum_{h=1}^{k} v_{kh}^2(l_k - l_h).$$

This shows that $\hat{l}_1 - l_1$ and $l_k - \hat{l}_k$ are essentially non-negative. Hence we need only prove that

$$(4.8)\qquad P(\hat{l}_1 - l_1 = 0) = 0; \qquad P(l_k - \hat{l}_k = 0) = 0.$$

We shall write down the proof for the first of these two equalities. The equation $F = ULU' = \Upsilon VLV'\Upsilon'$ (cf., (3.1a) and (3.2)) parametrizes (for this term see e.g., p. 246 of Busemann [1]) $\frac{1}{2}k(k+1)$-dimensional F-space (F is symmetric) in terms of the Cartesian product of $\frac{1}{2}k(k-1)$-dimensional V-space (V is orthogonal) and k-dimensional L-space. Now (4.7) shows that $\hat{l}_1 - l_1 = 0$ if and only if $l_1 = \cdots = l_g < l_{g+1}$ and $0 = v_{1,g+1} = \cdots = v_{1k} (g = 1, \cdots, k)$. Hence $\hat{l}_1 - l_1 = 0$ defines a union of subsets of the Cartesian product of V-space and L-space of dimension lower than $\frac{1}{2}k(k+1)$. In consequence the image of $\hat{l}_1 - l_1 = 0$ in F-space has $\frac{1}{2}k(k+1)$-dimensional Lebesgue measure zero, and as the probability distribution over F-space is assumed to be absolutely continuous with respect to this Lebesgue measure, it follows that $P(\hat{l}_1 - l_1 = 0) = 0$, (cf., equation (3.5) in Busemann [1]).

Theorem 2 is a trivial consequence of Lemma 1.

THEOREM 2. *Under the general assumption* A

$$(4.9) \qquad \mathcal{E}(l_1) < \lambda_1, \qquad \mathcal{E}(l_k) > \lambda_k, \qquad \mathcal{E}(l_k - l_1) > \lambda_k - \lambda_1.$$

PROOF. $\lambda_1 - \mathcal{E}(l_1) = \mathcal{E}(\hat{l}_1 - l_1) = \int (\hat{l}_1 - l_1)\, dP > 0$ by (3.4) and the first part of Lemma 1. A similar proof holds for the remaining two inequalities in (4.9).

Note that the amount of expectation-biases such as $\lambda_1 - \mathcal{E}(l_1)$, is a function of the distribution of F. If this distribution is specified, the expectation-biases can be calculated from formulae such as

$$(4.10a) \qquad \lambda_1 - \mathcal{E}(l_1) = \mathcal{E}(\hat{l}_1 - l_1) = \mathcal{E}[\sum_{h=1}^{k} v_{1h}^2 (l_h - l_1)]$$

(cf., (4.7)), which in case $k = 2$ simplifies into

$$(4.10b) \qquad \lambda_1 - \mathcal{E}(l_1) = \mathcal{E}[v_{12}^2 (l_2 - l_1)].$$

The next theorem will contain a result closely related to distribution-bias (cf., van der Vaart [7]; it is not really distribution-bias since $\hat{l}_1$ is not a statistic, see Section 3 of the present paper). Its proof depends on a simple lemma (see Section 4 in [7]), of which we shall cite a slightly altered version for easy reference (v without subscripts bears no relation to v with subscripts):

LEMMA 2. *Whether the random variables t and u are independent or dependent, if*

$$(4.11a) \qquad P(u > v) = 1,$$

then

$$(4.11b) \qquad P(t < \tau) \geqq P(t + u \leqq \tau + v).$$

A necessary and sufficient condition for equality of the two sides of (4.11*b*) *is*

$$(4.11c) \qquad P[(t + u > \tau + v) \cap (t < \tau)] = 0.$$

Under the weaker condition $P(u \geqq v) = 1$ either the first inequality sign in (4.11b) and the second one in (4.11c) should be replaced by $\geqq$, or the third inequality sign in (4.11b) should be replaced by $<$ and the first one in (4.11c) by $\geqq$.

We shall denote as Lemma 2′ the result obtained by reversing all six inequality signs in Lemma 2, except the second inequality sign in (4.11b).

THEOREM 3. *Under the general assumption* A

$$P(l_1 < \tau) \geqq P(\hat{l}_1 \leqq \tau) \text{ for any (real) } \tau, \tag{4.12a}$$

$$P(l_k > \tau) \geqq P(\hat{l}_k \geqq \tau) \text{ for any (real) } \tau, \tag{4.12b}$$

$$P(l_k - l_1 > \tau) \geqq P(\hat{l}_k - \hat{l}_1 \geqq \tau) \text{ for any (real) } \tau. \tag{4.12c}$$

Necessary and sufficient conditions for these inequalities to be strict are

$$P[(\hat{l}_1 > \tau) \cap (l_1 < \tau)] > 0, \tag{4.13a}$$

$$P[(\hat{l}_k < \tau) \cap (l_k > \tau)] > 0, \tag{4.13b}$$

$$P[(\hat{l}_k - \hat{l}_1 < \tau) \cap (l_k - l_1 > \tau)] > 0, \tag{4.13c}$$

respectively.

PROOF. Because of Lemma 1 we can apply Lemma 2. In Lemma 2 replace v by 0, t by l_1, u by $(\hat{l}_1 - l_1)$ to obtain (4.12a) and (4.13a). Again in Lemma 2′ replace v by 0, t by l_k, u by $(\hat{l}_k - l_k)$ to obtain (4.12b) and (4.13b), and v by 0, t by $(l_k - l_1)$, u by $(\hat{l}_k - l_k - \hat{l}_1 + l_1)$ to obtain (4.12c) and (4.13c).

The application of conditions (4.13) is easy if every subset of F-space, hence of the Cartesian product of V-space and L-space, which has a positive $\frac{1}{2}k(k+1)$-dimensional Lebesgue measure, has at the same time a positive probability (such is the case if F is distributed according to a nonsingular multinormal distribution): then all one has to show is that in (V-space) $\times$ (L-space) points exist in which both $\hat{l}_1 > \tau$ and $l_1 < \tau$, etc.; see pages 14 and 15 of van der Vaart [8].

THEOREM 4. *If the* $\frac{1}{2}k(k+1)$*-variate probability density function of the* $f_{ij}(1 \leqq i \leqq j \leqq k)$ *is symmetrical with respect to the point with coordinates* $\varphi_{ij} = \mathcal{E}(f_{ij})(1 \leqq i \leqq j \leqq k)$, then l_1, l_k, and $l_k - l_1$ *are negatively, positively and positively median-biased estimators of* λ_1, λ_k, *and* $\lambda_k - \lambda_1$, *respectively.*

PROOF. In inequalities (4.12) put $\tau = \lambda_1$, $\tau = \lambda_k$, $\tau = \lambda_k - \lambda_1$, respectively. Then proof will be complete if

$$P(\hat{l}_1 \leqq \lambda_1) = \tfrac{1}{2}, \qquad P(\hat{l}_k \geqq \lambda_k) = \tfrac{1}{2}, \qquad P(\hat{l}_k - \hat{l}_1 \geqq \lambda_k - \lambda_1) = \tfrac{1}{2}. \tag{4.14}$$

Now by (3.3), (3.2) and (3.1c) we have that

$$\hat{L} - \Lambda = \Upsilon' F \Upsilon - \Upsilon' \Phi \Upsilon = \Upsilon'(F - \Phi)\Upsilon.$$

Hence

$$\hat{l}_1 - \lambda_1 = v'_{\cdot 1}(F - \Phi)v_{\cdot 1}, \ \hat{l}_k - \lambda_k = v'_{\cdot k}(F - \Phi)v_{\cdot k}, \ \hat{l}_k - \hat{l}_1 - \lambda_k + \lambda_1$$
$$= \textstyle\sum_i \sum_j (v_{ik}v_{jk} - v'_{i1}v'_{j1})(f_{ij} - \varphi_{ij}).$$

These expressions when equated to zero represent hyperplanes in F-space which contain the center of symmetry. Hence (4.14) holds true.

The argument in the paragraph preceding Theorem 4 immediately yields a class of distributions, including the normal, of the matrix F for which median-bias is strict.

We want to emphasize the importance of the role of Lemma 1, i.e., essentially of assumption A, in our proofs. If, for instance, for the 2×2 matrix F the probability $P(f_{12} = 0) = 1$, and $P(f_{11} > c) = P(f_{22} < c) = 0$ (this example stems from T. W. Anderson in a personal communication (February 1958)), then $P(\hat{L} = L) = 1$ and all former deductions are invalid. Note that in this example our assumption A is violated: the distribution of F is not absolutely continuous relative to $\frac{1}{2}k(k+1)$ – dimensional Lebesgue measure.

Finally we shall give a few results concerning various moments of l_g and f_{ij}. We shall use the following equalities (where $M^2 = MM$ for any matrix M) in the proofs:

(4.15a) $$\operatorname{tr} \hat{L} = \operatorname{tr}(VLV') = \operatorname{tr} L = \operatorname{tr}(U'FU) = \operatorname{tr} F,$$

(4.15b) $$\operatorname{tr}(\hat{L}^2) = \operatorname{tr}(VLV'VLV') = \operatorname{tr}(VL^2V') = \operatorname{tr}(L^2) = \operatorname{tr}(U'FUU'FU) = \operatorname{tr}(U'F^2U) = \operatorname{tr}(F^2),$$

(4.15c) $$\operatorname{tr}(\Phi^2) = \operatorname{tr}(\Lambda^2) = \operatorname{tr}[\mathcal{E}(\hat{L})\cdot\mathcal{E}(\hat{L})].$$

Proofs of these equalities are easy: apply (4.2), (3.3), (3.1c) and (3.4).

THEOREM 5. *Under the general assumption* A *we have for the sum of mean square errors*

(4.16) $$\sum_g \mathcal{E}(l_g^2 - \lambda_g^2) = \sum_{g,h} \operatorname{var} \hat{l}_{gh} = \sum_{i,j} \operatorname{var} f_{ij};$$

for the sum of variances

(4.17a) $$\sum_g \operatorname{var} l_g = \sum_{i,j} \operatorname{var} f_{ij} + \sum_g \lambda_g^2 - \sum_g (\mathcal{E} l_g)^2$$

(4.17b) $$= \sum_{i,j} \operatorname{var} f_{ij} + \sum_{i,j} \varphi_{ij}^2 - \sum_g (\mathcal{E} l_g)^2;$$

for the sum of covariances

(4.18) $$\sum_{g,h} \operatorname{cov}(l_g, l_h) = \sum_{g,h} \operatorname{cov}(\hat{l}_g, \hat{l}_h) = \sum_{i,j} \operatorname{cov}(f_{ii}, f_{jj}).$$

Here $\sum_{i,j}$ stand for $\sum_{i=1}^k \sum_{j=1}^k$, $\sum_g$ for $\sum_{g=1}^k$.

PROOF. Because of (4.15b) and (4.15c)

$$\sum_{i,j} \operatorname{var} f_{ij} = \sum_{i,j} \mathcal{E}(f_{ij}^2) - \sum_{i,j} (\mathcal{E} f_{ij})^2 = \mathcal{E} \operatorname{tr}(F^2) - \operatorname{tr}(\Phi^2) = \mathcal{E} \operatorname{tr}(\hat{L}^2) - \operatorname{tr}[\mathcal{E}(\hat{L})\cdot\mathcal{E}(\hat{L})] = \sum_{g,h} \operatorname{var} \hat{l}_{gh} = \mathcal{E} \operatorname{tr}(L^2) - \operatorname{tr}(\Lambda^2) = \sum_g \mathcal{E}(l_g^2 - \lambda_g^2).$$

Equation (4.17a) follows immediately from (4.16); (4.15c) then shows the equivalence of (4.17a) and (4.17b). Finally apply (4.15a) to prove

$$\sum_{g,h} \operatorname{cov}(l_g, l_h) = \mathcal{E}(\operatorname{tr} L)^2 - (\mathcal{E} \operatorname{tr} L)^2 = \mathcal{E}(\operatorname{tr} \hat{L})^2 - (\mathcal{E} \operatorname{tr} \hat{L})^2 = \mathcal{E}(\operatorname{tr} F)^2 - (\mathcal{E} \operatorname{tr} F)^2,$$

which serves to prove (4.18).

As for the consequence of this theorem concerning response surface theory see the discussion round equation (14) in van der Vaart [9]. Here we repeat only that if $k = 2$ equation (4.17a) yields

$$\sum_g \operatorname{var} l_g = \sum_{i,j} \operatorname{var} f_{ij} - 2\alpha_\beta^2 - 2(\lambda_2 - \lambda_1)\alpha_\beta , \tag{4.19}$$

where $\alpha_\beta = \mathcal{E}(l_2 - \lambda_2) = -\mathcal{E}(l_1 - \lambda_1) > 0$. This equation is important in theoretical work: in theoretical investigations the data of a problem frequently are λ_1, λ_2, $\operatorname{var} f_{ij}$. From (4.19) it appears that in cases where $\operatorname{var} l_1 = \operatorname{var} l_2$, the knowledge of the first order moment α_β is sufficient to calculate the second order moment $\operatorname{var} l_g$. Hence the problem arises to find conditions under which $\operatorname{var} l_1 = \operatorname{var} l_2 (k = 2)$. We replace this problem by the following: to find conditions under which a kind of symmetry exists in the distribution of L such that $\operatorname{var} l_1 = \operatorname{var} l_k$.

Let q denote the joint probability density of $l_1, \cdots, l_k$. A type of symmetry which suits our purpose is defined by

$$q(l_1, l_2, \cdots, l_k) = q(\gamma - l_k, \gamma - l_{k-1}, \cdots, \gamma - l_1). \tag{4.20}$$

For, in the first place, because of (2.5) q should be zero except for

(a) $l_1 \leqq l_2 \leqq \cdots \leqq l_k$; (b) $\gamma - l_k \leqq \gamma - l_{k-1} \leqq \cdots \leqq \gamma - l_1$.

Conditions (a) and (b) coincide. In order to show other useful features of a distribution of L satisfying (4.20) we introduce the matrix

$$^*L = \| l_{k+1-h}\delta_{gh} \| \begin{matrix} g = 1 \cdots k \\ h = 1 \cdots k \end{matrix}, \tag{4.21}$$

cf., the definition of L and Λ in (2.6). Now the best way to describe the type of symmetry determined by (4.20) is by means of

$$L \sim \gamma I - {}^*L \tag{4.22}$$

(I is the identity matrix; the sign $\sim$ was defined in (2.3a)). As the expectations of functions of (stochastically) equivalent matrices are equal (cf., (2.3b)) we find from (4.22) that $\mathcal{E}(L) + \mathcal{E}(^*L) = \gamma I$, whence

$$\mathcal{E}l_g + \mathcal{E}l_{k+1-g} = \gamma \qquad (g = 1 \cdots k). \tag{4.23a}$$

From the definition of $^*L$, (4.23a) and Theorem 1 we find

$$2\,\mathcal{E} \operatorname{tr} L = 2\,\mathcal{E} \operatorname{tr} {}^*L = 2 \operatorname{tr} \Lambda = 2 \operatorname{tr} \Phi = k\gamma. \tag{4.23b}$$

Now consider the direct product of L and L (cf., p. 81 of MacDuffee [5]); as $L = L$, we need not distinguish between left and right direct product. Evidently

$$L \otimes L \sim (\gamma I - {}^*L) \otimes (\gamma I - {}^*L).$$

Taking expectations and subtracting

$$(\mathcal{E}L) \otimes (\mathcal{E}L) = [\mathcal{E}(\gamma I - {}^*L)] \otimes [\mathcal{E}(\gamma I - {}^*L)]$$

we find

$$\mathcal{E}[L \otimes L - (\mathcal{E}L) \otimes (\mathcal{E}L)] = \mathcal{E}[{}^*L \otimes {}^*L - (\mathcal{E}{}^*L) \otimes (\mathcal{E}{}^*L)],$$

whence

$$\text{(4.23c)} \qquad \text{cov}\,(l_{g_1}, l_{g_2}) = \text{cov}\,(l_{k+1-g_1}, l_{k+1-g_2}) \quad (g_1, g_2 = 1 \cdots k).$$

so that for $g_1 = g_2 = g$

$$\text{(4.23d)} \qquad \text{var}\, l_g = \text{var}\, l_{k+1-g} \qquad (g = 1 \cdots k),$$

i.e., the equality, which prompted this part of our investigation. We shall now give a condition on the distribution of F, sufficient to ensure the equivalence of L and $\gamma I - {}^*L$. This condition will ensure *a fortiori* that (4.23d) holds true.

THEOREM 6. *Let F be a real symmetric random $k \times k$ matrix with $\mathcal{E}F = \Phi$. Then in order that $L \sim \gamma I - {}^*L$ (L and ${}^*L$ consisting of the latent roots of F according to (2.6) and (4.21)), it is sufficient that some pair of real, non-singular, random or non-random, $k \times k$ matrices, M_1 and M_2, exists such that*

$$\text{(4.24)} \qquad M_1^{-1}FM_1 \sim \gamma I - M_2^{-1}FM_2 .$$

This condition entails that

$$\text{(4.25a)} \qquad 2\,\text{tr}\,\mathcal{E}(F) = 2\,\text{tr}\,\Phi = k\gamma$$

and in case M_1 and M_2 are not random

$$\text{(4.25b)} \qquad \sum_{i,j} \varphi_{ij}[\{M_1^{-1}\}_{gi}\{M_1\}_{jh} + \{M_2^{-1}\}_{gi}\{M_2\}_{jh}] = \gamma \cdot \delta_{gh} \quad (g, h = 1 \cdots k)$$

PROOF. Denote the functional relation which to any matrix assigns the matrix of its latent roots (ordered according to increasing magnitude) by Ψ; then $L = \Psi(F)$. It is well known that under the conditions of the theorem

$$\text{(4.26a)} \qquad \Psi(M_1^{-1}FM_1) = \Psi(F) = L.$$

Likewise, since $\gamma - l_k \leqq \gamma - l_{k-1} \leqq \cdots \leqq \gamma - l_1$ are the latent roots of $\gamma I - F$ if $l_1 \leqq \cdots \leqq l_k$ are the latent roots of F: $\Psi(\gamma I - F) = \gamma I - {}^*L$, whence

$$\text{(4.26b)} \qquad \Psi(\gamma I - M_1^{-1}FM_2) = \Psi(\gamma I - F) = \gamma I - {}^*L.$$

Since Ψ is continuous, comparison of (4.26a) and (4.26b) proves the first part of the theorem. Equation (4.25a) coincides with (4.23b) and may, of course, be proved directly from (4.24). Equation (4.25b) holds because the expectations of corresponding elements of equivalent matrices are equal (the symbol $\{A\}_{ij}$, A any matrix, stands for a_{ij}).

In conclusion we want to give an application of Theorem 6. Take $k = 2$, $M_1 = I$, $M_2 = \begin{Vmatrix} 0 & -1 \\ 1 & 0 \end{Vmatrix}$. Then $M_2^{-1}F\,M_2 = \begin{Vmatrix} f_{22} & -f_{21} \\ -f_{12} & f_{11} \end{Vmatrix} = {}^*F$, say. According to Theorem 6 we have that $L \sim \gamma I - {}^*L$ if

$$\text{(4.27)} \qquad F \sim \gamma I - {}^*F.$$

This equivalence has some interesting consequences with respect to the first and second order moments of the distribution of F. By an argument exactly like the argument leading to the various equations (4.23) we find $\mathcal{E}(F) + \mathcal{E}(^*F) = \gamma I$, $\mathcal{E}[F \otimes F - (\mathcal{E}F) \otimes (\mathcal{E}F)] = \mathcal{E}[^*F \otimes {}^*F - (\mathcal{E}^*F) \otimes (\mathcal{E}^*F)]$, whence

$$\mathcal{E}f_{11} + \mathcal{E}f_{22} = \gamma, \tag{4.27a}$$

$$\text{var}\, f_{11} = \text{var}\, f_{22}, \qquad \text{cov}\,(f_{11}, f_{12}) = -\text{cov}\,(f_{12}, f_{22}). \tag{4.27b}$$

There are no consequences of (4.27) with respect to $\mathcal{E}f_{12}$, var f_{12}, cov (f_{11}, f_{22}). As a corollary, if F is normally distributed, (4.27a) and (4.27b) are sufficient in order that $q(l_1, l_2) = q(\gamma - l_2, \gamma - l_1)$, whence var l_1 = var l_2.

Acknowledgment. My thanks are due to Professor Dr. R. J. Hader of the Institute of Statistics, North Carolina State College of the University of North Carolina, Raleigh, for his suggestion that it might be interesting to investigate some questions in the theory of response surface estimation.

REFERENCES

[1] Herbert Busemann, "Intrinsic area," *Ann. Math.*, Vol. 48, (1947), pp. 234–267.
[2] J. L. Doob, *Stochastic Processes*, John Wiley and Sons, New York, 1953.
[3] F. R. Gantmacher, *Matrizenrechnung*, Teil 1, *Allgemeine Theorie*, Deutscher Verlag der Wissenschaften, Berlin, 1958.
[4] A. N. Kolmogorov, *Foundations of the Theory of Probability*, English transl. ed. by N. Morrison, Chelsea Pub. Co., New York, 1950.
[5] C. C. MacDuffee, *The Theory of Matrices*, corrected reprint, Chelsea Pub. Co., New York, 1946.
[6] Stanisław Saks, *Theory of the Integral*, Monografie Matematyczne, Vol. 7, 2nd rev. ed., transl. L. C. Young, Hafner, New York, 1937.
[7] H. R. van der Vaart, "Some extensions of the idea of bias", *Ann. Math. Stat.*, Vol. 32 (1961), pp. 436–447.
[8] H. R. van der Vaart, "Some results on the probability distribution of the latent roots of a symmetric matrix of continuously distributed elements, and some applications to the theory of response surface estimation," Institute of Statistics Mimeo Series No. 189, Raleigh, 1958.
[9] H. R. van der Vaart, "On certain types of bias in current methods of response surface estimation," *Bull. Inst. Internat. Stat.*, Vol. 37, No. 3 (1960), pp. 191–203.