# On Symmetric, Orthogonal, and Skew-Symmetric Matrices

By P. L. Hsu

(*Received* 9*th July* 1948. *Read* 5*th November* 1948.)

1. Introduction and Notation. In this paper all the scalars are real and all matrices are, if not stated to be otherwise, $p$-rowed square matrices. The diagonal and superdiagonal elements of a symmetric matrix, and the superdiagonal elements of a skew-symmetric matrix, will be called the distinct elements of the respective matrices. $\Sigma$ will denote both the set of all symmetric matrices and the $\frac{1}{2}p(p+1)$-dimensional space whose coordinates are the distinct elements arranged in some specific order. $K$ will denote both the set of all skew-symmetric matrices and the $\frac{1}{2}p(p-1)$-dimensional space whose coordinates are the distinct elements arranged in some specific order. Any sub-set of $\Sigma$ ($K$) will mean both the sub-set of symmetric (skew-symmetric) matrices and the set of points of $\Sigma$ ($K$). Any point function defined in $\Sigma$ ($K$) will be written as a function of a symmetric (skew-symmetric) matrix. $D_a$ will denote the diagonal matrix whose diagonal elements are $a_1, a_2, \ldots, a_p$. The characteristic roots of a symmetric matrix will be called its roots.

The orthogonal matrices $\Gamma$ with $|\Gamma + I| \neq 0$ and the skew-symmetric matrices $X$ are in (1, 1)-correspondence, on account of the following pair of equivalent equations:

(1) $$\Gamma = 2(I+X)^{-1} - I,$$

(1′) $$X = 2(I+\Gamma)^{-1} - I.$$

That the skew-symmetry of $X$ implies the orthogonality of $\Gamma$, and vice versa,[1] is the direct result of the following computation:

$$\begin{aligned}[2(I+X)^{-1} - I]\ [2(I+X)^{-1} - I]' \\ &= (I+X)^{-1}\ (I-X)\ \ (I+X)(I-X)^{-1} \\ &= (I+X)^{-1}(I+X)(I-X)(I-X)^{-1} = I,\end{aligned}$$

$$\begin{aligned}&[2(I+\Gamma)^{-1} - I] + [2(I+\Gamma)^{-1} - I]' \\ &= 2(I+\Gamma)^{-1} + 2(I+\Gamma)^{-1}\Gamma - 2I = 2(I+\Gamma)^{-1}(I+\Gamma) - 2I = 0.\end{aligned}$$

If $S \in \Sigma$, and if its roots are $\theta_1, \ldots, \theta_p$, it is well known that

(2) $$S = \Delta D_\theta \Delta'$$

where $\Delta$ is some orthogonal matrix.

---

[1] This result, together with a lengthy derivation, is given in Kowalewski, *Einführung in die Determinantentheorie* (Leipzig, 1909), pp. 171-175.

LEMMA 1. *If $A$ is any $p \times p$ matrix, there is a matrix $D_\epsilon$, with $\epsilon_i = +1$ or $-1$ $(i = 1, \ldots, p)$, such that* $| A + D_\epsilon | \neq 0$.

*Proof.* The lemma is evidently true for $p = 1$. Assume it is true for $p - 1$. Write

$$A = \begin{pmatrix} A_1 & b' \\ c & d \end{pmatrix},$$

where $A_1$ is $(p-1) \times (p-1)$ and $b$, $c$ are rows. By assumption there is a $D_\eta$, with $\eta_i = +1$ or $-1$ $(i = 1, \ldots, p-1)$, such that $| A_1 + D_\eta | \neq 0$. Since

$$(3) \qquad \begin{vmatrix} A_1 + D_\eta & b' \\ c & d+1 \end{vmatrix} - \begin{vmatrix} A_1 + D_\eta & b' \\ c & d-1 \end{vmatrix} = 2 \, | A_1 + D_\eta | \neq 0,$$

the two determinants on the left side of (3) cannot both vanish. Hence the lemma is proved.

LEMMA 2. *If $S \in \Sigma$, there is an orthogonal $\Gamma$ with $| \Gamma + I | \neq 0$ such that $S = \Gamma D_\theta \Gamma'$.*

*Proof.* For the $\Delta$ in (2) there is, by Lemma 1, a $D_\epsilon$ with $| \Delta + D_\epsilon | \neq 0$. Hence $| \Delta D_\epsilon + I | \neq 0$. Moreover, $\Delta D_\epsilon$ also satisfies (2). Hence the matrix $\Gamma = \Delta D_\epsilon$ answers all the requirements.

Combining Lemma 2 with (1) we obtain

THEOREM 1. *If $S \in \Sigma$, there is an $X \in K$ such that*

$$(4) \qquad S = [2(I + X)^{-1} - I] \, D_\theta \, [2(I + X)^{-1} - I]'.$$

In § 2 we investigate the uniqueness of the expression (4) for a given $S$. (4) expresses each distinct element of $S$ as a function of $\frac{1}{2}p(p+1)$ arguments, viz., the distinct elements of $X$ and the $\theta$'s. In § 3 we evaluate the functional determinant of these functions or, to use another term, the Jacobian of the transformation (4). In § 4 we apply the results to prove a formula for the integral of some function of $S$.

2. The symmetric matrices $S$ which have multiple roots satisfy the equation $f(S) = 0$, where $f(S)$ is the discriminant of the characteristic equation of $S$. Let $F$ be the surface $f(S) = 0$. If $S \in \Sigma - F$,

we shall so name the roots $\theta_1, \ldots, \theta_p$ that $\theta_1 > \theta_2 > \ldots > \theta_p$. Thus each $\theta_i$ is a well-defined function of $S$. The equation

(5) $$S = \Gamma D_\theta \Gamma',$$

where $\Gamma = (\gamma_{ij})$ is orthogonal, then uniquely determines every column of $\Gamma$ except for a sign. In particular,

$$|\gamma_{1i}| = g_i(S) \qquad (i=2,\ldots,p)$$

are well-defined functions of $S$. Let $F_i$ be the surface $g_i(S) = 0$. If $S \in \Sigma - F - F_2 - \ldots - F_p$, we take $\gamma_{1i} = -g_i(S)$ $(i = 2, \ldots, p)$. Thus all the columns except the first one of $\Gamma$ are uniquely determined, since the signs of the top elements of these columns are determined. Hence we have

LEMMA 3. *If $S \in \Sigma - F - F_2 - \ldots - F_p$, equation* (5) *has exactly two solutions for $\Gamma$: $\Gamma_1$ and $\Gamma_2$, whose elements* (1, 2), (1, 3), ..., (1, $p$) *are all negative. $\Gamma_1$ and $\Gamma_2$ differ only in the respect that the first column of one is the negative of the first column of the other.*

By virtue of Lemma 3 all the elements $\gamma_{ij}$ $(i, j = 2, \ldots, p)$ of $\Gamma$ are well-defined functions of $S$. Writing $\Gamma^{(1)} = (\gamma_{ij})_{(i,\, j=2,\ldots,p)}$ we have $|\Gamma^{(1)} + I| = h(S)$, a well-defined function of $S$. Let $F'$ be the surface $h(S) = 0$. Let $E = \Sigma - F - F_2 - \ldots - F_p - F'$.

LEMMA 4. *If $S \in E$, one of the determinants $|\Gamma_1 + I|$ and $|\Gamma_2 + I|$ in Lemma* 3 *is zero while the other is not zero.*

*Proof.* By Lemma 3 and the defiinition of $E$ we have

$$|\Gamma_1 + I| + |\Gamma_2 + I| = 2\,|\Gamma^{(1)} + I| \neq 0.$$

Hence one of the determinants is not zero. Suppose $|\Gamma_1 + I| \neq 0$. Then, by (1), $\Gamma_1 = 2(I + X)^{-1} - I$, where $X$ is skew-symmetric. Letting $H$ be the diagonal matrix $[-1, 1, \ldots, 1]$ we have $\Gamma_2 = \Gamma_1 H$. Hence

$$\begin{aligned}|\Gamma_2 + I| = |\Gamma_1 H + I| &= |2(I + X)^{-1} H + I - H| \\ &= |X + I|^{-1}\,|2H + (I + X)(I - H)|.\end{aligned}$$

It is easily seen that the first row of the matrix $2H + (I + X)(I - H)$ consists of zeros. Hence $|\Gamma_2 + I| = 0$. This completes the proof.

If we take $\Gamma$ to be $\Gamma_1$ or $\Gamma_2$ according as $|\Gamma_1 + I| \neq 0$ or $|\Gamma_2 + I| \neq 0$ we obtain

LEMMA 5. *If $S \in E$, there is a unique $\Gamma$ which satisfies* (5) *and the following conditions:*

https://doi.org/10.1017/S001309150001422X Published online by Cambridge University Press

$$\text{(i)} \quad \gamma_{1i} < 0 \qquad (i = 2, \ldots, p)$$

$$\text{(ii)} \quad | \Gamma + I | \neq 0.$$

By combining Lemma 5 with (1) and noticing that in (1) the non-diagonal elements of $\Gamma$ and $(I + X)^{-1}$ must have the same sign, we obtain

THEOREM 2. *If $S \epsilon E$, the equation* (4) *has a unique solution in $X$ such that the elements* (1, 2), (1, 3), ..., (1, $p$) *of* $(I + X)^{-1}$ *are all negative.*

Let $M$ be the sub-set of $K$ defined by the condition that the elements (1, 2), (1, 3), ..., (1, $p$) of $(I + X)^{-1}$ are all negative, let $\Theta$ be the sub-set of the space of $(\theta_1, \ldots, \theta_p)$ defined by the condition $\theta_1 > \theta_2 > \ldots > \theta_p$, and let $E^* = M \times \Theta$, a sub-set of the $\frac{1}{2}p(p+1)$-dimensional space. Theorem 2 asserts that the equation (4) effects a (1, 1)-mapping of $E^*$ on $E$. Notice also that by the definition of $E$, it differs from $\Sigma$ by a set of measure zero.

3. In order to facilitate the computation of the functional determinant[1] for (4) we shall adopt the following notation. If $x = (x_1, \ldots, x_m)$ and $y = (y_1, \ldots, y_m)$ and if

$$x_i = \sum_{j=1}^{m} a_{ij} y_j \qquad (i = 1, \ldots, m),$$

then we define the symbol $D(x; y)$ to be the discriminant $| a_{ij} |$.

LEMMA 6.

$$\text{(i)} \quad D(x; y) = \frac{1}{D(y; x)}$$

$$\text{(ii)} \quad D(x; y) = D(x; z)\, D(z; y).$$

*Proof.* (i) is an immediate consequence of definition; (ii) is a special case of the multiplicative law of functional determinants.

The extension of (ii) to more than two factors is obvious.

In the equation

$$X = AYA', \tag{6}$$

$X$ is symmetric or skew-symmetric according as $Y$ is symmetric or skew-symmetric. In either case (6) expresses each distinct element of $X$ as a linear function of the distinct elements of $Y$, with coefficients depending on $A$. The following lemma gives the discriminant $D(X; Y)$ of this set of linear functions.

[1] Functional determinants are considered as determined up to a sign. In all the computations in this section signs of functional determinants are neglected.

LEMMA 7. *For* (6) *we have*

$$D(X;\,Y) = \begin{cases} |A|^{p+1} & \text{if } Y \text{ is symmetric} \\ |A|^{p-1} & \text{if } Y \text{ is skew-symmetric.} \end{cases} \tag{7}$$

*Proof.* It is sufficient to prove (7) for a non-singular $A$. Now every non-singular matrix is a product of a finite number of matrices of either of the following two types, (a) a diagonal matrix whose diagonal elements are 1 with the exception of one, which is, say, $a$; (b) a matrix whose diagonal elements are 1 and whose non-diagonal elements are 0 with the exception of one. Further, if we denote $D(X;\,Y)$ for (6) by $P(A)$, a moment of reflection will show that $P(AB) = P(A)\,P(B)$. Hence it is sufficient to prove (7) for $A$ belonging to either of the types (a) and (b), i.e. $D(X;\,Y) = \begin{cases} a^{p+1} \\ a^{p-1} \end{cases}$ if $A$ is of type (a) and $D(X;\,Y) = 1$ for both cases if $A$ is of type (b). The proof of these assertions is easy and is left to the reader.

If $x_1, \ldots, x_m$ are functions of $y_1, \ldots, y_m$, then by definition the functional determinant is equal to $D(dx;\,dy)$, where $dx$ denotes the system $dx_1, \ldots, dx_m$. For brevity we write $dA$ for the matrix whose elements are the differentials of the elements of $A$.

In order to differentiate (1) we use the formula

$$dA^{-1} = -A^{-1}(dA)A^{-1}.$$

Then

$$d\Gamma = -2(I+X)^{-1}dX(I+X)^{-1} = -\tfrac{1}{2}(I+\Gamma)\,dX\,(I+\Gamma). \tag{8}$$

Let $J$ be the functional determinant for (4), i.e. for

$$S = \Gamma D_\theta \Gamma' \tag{9}$$

where $\Gamma$ is given by (1). Differentiating (9) and using (8) we have

$$dS = \Gamma D_\theta d\Gamma' + (d\Gamma)\,D_\theta\Gamma' + \Gamma D_\theta\Gamma' = \tfrac{1}{2}\Gamma D_\theta(I+\Gamma')\,dX\,(I+\Gamma') - \tfrac{1}{2}(I+\Gamma)\,dX\,(I+\Gamma)\,D_\theta\Gamma' + \Gamma D_{d\theta}\Gamma'. \tag{10}$$

Also,

$$J = D(dS;\,dX,\,d\theta). \tag{11}$$

From (10) we get

$$\Gamma' dS\,\Gamma = \tfrac{1}{2}D_\theta(I+\Gamma')\,dX\,(I+\Gamma) - \tfrac{1}{2}(I+\Gamma')\,dX\,(I+\Gamma)\,D_\theta + D_{d\theta}. \tag{12}$$

If we write $\xi_i$ for $d\theta_i$, and define $U$ and $Y$ by

$$U = \Gamma' dS\,\Gamma \tag{13}$$

$$Y = (I+\Gamma')\,dX\,(I+\Gamma), \tag{14}$$

then (12) gives

$$U = \tfrac{1}{2} D_\theta Y - \tfrac{1}{2} Y D_\theta + D_\xi,$$

i.e.,

(15) $$u_{ii} = \xi_i \qquad (i = 1, \ldots, p),$$
$$u_{ij} = \tfrac{1}{2}(\theta_i - \theta_j) y_{ij} \qquad (i < j).$$

(13), (14) and (15) give respectively

$$D(dS;\, U) = \frac{1}{D(U;\, dS)} = 1, \text{ by Lemmas 6 and 7;}$$

$$D(Y;\, dX) = 2^{p(p-1)} \mid I + X \mid^{-(p-1)}, \text{ by Lemma 6 and (1);}$$

$$D(U;\, Y, \xi) = 2^{-\frac{1}{2}p(p-1)} \prod_{i<j} (\theta_i - \theta_j).$$

Hence, by (11) and Lemma 7,

$$J = D(dS;\, U)\, D(U;\, Y, \xi)\, D(Y, \xi;\, dX, \xi)$$
$$= D(dS;\, U)\, D(U;\, Y, \xi)\, D(Y;\, dX),$$

and so, finally,

(16) $$J = 2^{\frac{1}{2}p(p-1)} \mid I + X \mid^{-(p-1)} \prod_{i<j} (\theta_i - \theta_j).$$

4. Let us first evaluate the following integrals:

(17) $$A_p = \int_K \mid I + X \mid^{-(p-1)} dm$$

(18) $$B_p = \int_M \mid I + X \mid^{-(p-1)} dm$$

where $X$ is a $p$-rowed skew-symmetric matrix, $dm$ denotes the volume element, $K$ is the $\frac{1}{2}p(p-1)$-dimensional space, and $M$ is defined in the last paragraph of § 2.

We write

$$X = \begin{pmatrix} 0 & x \\ -x' & Y \end{pmatrix}, \qquad x = (x_1, \ldots, x_{p-1})$$

so that

$$\mid I + X \mid = \begin{vmatrix} 1 & X \\ -x' & I + Y \end{vmatrix} = \mid I + Y \mid (1 + x(I+Y)^{-1}x').$$

Also, the elements $(1, 2), (1, 3), \ldots, (1, p)$ of $(I+X)^{-1}$ are the elements of the row $-(1 + x(I+Y)^{-1}x')^{-1} x (I+Y)^{-1}$. Hence $M$ is the set of points such that every element of the row $x(I+Y)^{-1}$ is positive. Making the transformation

$$x' = (I + Y)\, u', \qquad u = (u_1, \ldots, u_{p-1})$$

whose Jacobian is $\mid I + Y \mid$, we obtain

https://doi.org/10.1017/S001309150001422X Published online by Cambridge University Press

$$x(I+Y)^{-1}x' = u(I+Y)u' = uu' = u_1^2 + \ldots + u_{p-1}^2.$$

Hence

$$(19)\quad A_p = \int_{K_1} |I+Y|^{-(p-2)} dm \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} (1 + u_1^2 + \ldots + u_{p-1}^2)^{-(p-1)} du_1 \ldots du_{p-1},$$

$$(20)\quad B_p = \int_{K_1} |I+Y|^{-(p-2)} dm \int_0^{\infty} \ldots \int_0^{\infty} (1 + u_1^2 + \ldots + u_{p-1}^2)^{-(p-1)} du_1 \ldots du_{p-1},$$

where $K_1$ is the $\frac{1}{2}(p-1)(p-2)$-dimensional space. It follows from (19) and (20) that

$$(21)\qquad B_p = 2^{-(p-1)} A_p,$$

$$(22)\quad A_p = A_{p-1} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} (1 + u_1^2 + \ldots + u_{p-1}^2)^{-(p-1)} du_1 \ldots du_{p-1}$$
$$= \frac{\pi^{p/2}}{2^{p-2}\,\Gamma(p/2)} A_{p-1},$$

and the easy computation $A_2 = \pi$ leads to the result

$$(23)\qquad A_p = \pi \prod_{r=3}^{p} \frac{\pi^{r/2}}{2^{r-2}\,\Gamma(r/2)} = \frac{\pi^{p(p+1)/4}}{2^{(p-1)(p-2)/2} \prod_{r=1}^{p} \Gamma(r/2)},$$

whence also

$$(24)\qquad B_p = \frac{\pi^{p(p+1)/4}}{2^{p(p-1)/2} \prod_{r=1}^{p} \Gamma(r/2)}.$$

We can now prove

THEOREM 3. *Let $f(S)$ be a function of the distinct elements of the $p$-rowed symmetric matrix $S$, such that $f(S)$ is a function only of the roots $\theta_1, \ldots, \theta_p$ of $S$, $\theta_1 \geqq \theta_2 \geqq \ldots \geqq \theta_p$:*

$$f(S) = g(\theta_1, \ldots, \theta_p).$$

*Let $\Sigma$ be the $\frac{1}{2}p(p+1)$-dimensional space. Then*

$$(25)\quad \int_{\Sigma} f(S)\,dm = \frac{\pi^{p(p+1)/4}}{\prod_{r=1}^{p} \Gamma(r/2)} \int_{\Theta} g(\theta_1, \ldots, \theta_p) \prod_{i<j} (\theta_i - \theta_j)\, d\theta_1 \ldots d\theta_p,$$

*where $\Theta$ is the domain $\theta_1 > \theta_2 > \ldots > \theta_p$.*

*Proof.* By the remark at the end of §2 we have

$$\int_{\Sigma} f(S)\,dm = \int_{E} f(S)\,dm.$$

https://doi.org/10.1017/S001309150001422X Published online by Cambridge University Press

By Theorem 2 we may use the transformation (4), with the Jacobian (16). Hence

$$\int_{\Sigma} f(S)\, dm = 2^{\frac{1}{2}p(p-1)} \int_{E^*} | I + X |^{-(p-1)} g(\theta_1, \ldots, \theta_p) \prod_{i<j} (\theta_i - \theta_j)\, dm$$

where $E^* = M \times \Theta$. Hence

$$\int_{\Sigma} f(S)\, dm = 2^{\frac{1}{2}p(p-1)} B_p \int_{\Theta} g(\theta_1, \ldots, \theta_p) \prod_{i<} (\theta_i - \theta_j)\, d\theta_1, \ldots d\theta_p.$$

Using (24) we get (25).

Theorem 3 implies the following theorem on the probability distribution of roots, which is an important subject in statistics.

Theorem 4. *Let the distinct elements of the p-rowed symmetric matrix S be random variables whose joint distribution in $\Sigma$ has a probability density $f(S)$, such that $f(S)$ is a function only of the roots $\theta_1 \geqq \ldots \geqq \theta_p$ of S:*

$$f(S) = g(\theta_1, \ldots, \theta_p).$$

*Then the joint distribution of roots in $\Theta$ has the probability density function*

$$\frac{\pi^{p(p+1)/4}}{\prod_{r=1}^{p} \Gamma(r/2)} g(\theta_1, \ldots, \theta_p) \prod_{i<j} (\theta_i - \theta_j).$$

*Proof.* If $A$ is any sub-set of $\Theta$, let $B$ be the sub-set of $\Sigma$ such that $(\theta_1, \ldots, \theta_p) \in A$. Then

$$\Pr\left\{ \theta_1, \ldots, \theta_p) \in A \right\} = \int_{\Sigma} \phi(S) f(S)\, dm, \tag{26}$$

where $\phi(S) = 1$ or $0$ according as $S \in B$ or $S \bar{\in} B$. Now $\phi(S) = \psi(\theta_1, \ldots, \theta_p)$, where $\psi = 1$ or $0$ according as $(\theta_1, \ldots, \theta_p) \in A$ or not. Hence $\phi(S) f(S) = \psi(\theta_1, \ldots, \theta_p) g(\theta_1, \ldots, \theta_p)$. Applying Theorem 3 to (26) we get

$$\Pr\{(\theta_1, \ldots, \theta_p) \in A\}$$

$$= \frac{\pi^{p(p+1)/4}}{\prod_{r=1}^{p} \Gamma(r/2)} \int_{\Theta} \psi(\theta_1, \ldots, \theta_p) g(\theta_1, \ldots, \theta_p) \prod_{i<j} (\theta_i - \theta_j)\, d\theta_1 \ldots d\theta_p$$

$$= \frac{\pi^{p(p+1)/4}}{\prod_{r=1}^{p} \Gamma(r/2)} \int_{A} g(\theta_1, \ldots, \theta_p) \prod_{i<j} (\theta_i - \theta_j)\, d\theta_1 \ldots d\theta_p,$$

which proves the theorem.

University of North Carolina,
U.S.A.