THE INVERSE OF A CERTAIN MATRIX, WITH AN APPLICATION[1]

By Chandan K. Mustafi

Columbia University

Let $\mathbf{I}_m$ be the $m \times m$ identity matrix and $\mathbf{W}_m^{(k)}$ be an $m \times m$ matrix whose upper left $k \times k$ submatrix consists of elements equal to one and remaining elements equal to zero. In this note, we consider the problem of finding the inverse of the matrix

$$\mathbf{I}_m + \sum_{k=1}^m c_k \mathbf{W}_m^{(k)},$$

with $0 < c_k < \infty$ $(k = 1, 2, \cdots, m)$.

The special case $c_1 = c_2 = \cdots = c_m$ will be discussed in some detail. An application of the result will also be pointed out.

Theorem. $$(\mathbf{I}_m + \sum_{k=1}^m c_k \mathbf{W}_m^{(k)})^{-1} = \mathbf{I}_m - \mathbf{Q},$$

where $\mathbf{Q} = (q_{ij})$ *is an* $m \times m$ *matrix whose elements* q_{ij} *have the following form:*

$$(1) \qquad \begin{aligned} q_{ij} &= \lambda_i q_j, \; i \leqq j, \\ &= \lambda_j q_i, \; i > j, \qquad (i, j = 1, 2, \cdots, m). \end{aligned}$$

The quantities $\lambda_3, \cdots, \lambda_m$ *and* $q_1, q_2, \cdots, q_m$ *are obtained from the following equations:*

$$(2) \qquad c_{r-1}\lambda_{r+1} - \lambda_r(c_{r-1} + c_r + c_{r-1}c_r) + \lambda_{r-1}c_r = 0 \quad (r = 2, 3, \cdots, m-1),$$

with $\lambda_1 = 1, \lambda_2 = 1 + c_1$;

$$(3) \qquad \lambda_{r-1}q_r - \lambda_r q_{r-1} + c_{r-1} = 0 \qquad (r = 2, 3, \cdots, m);$$

$$(4) \qquad q_m = c_{m-1}c_m[c_{m-1}(1 + c_m)\lambda_m + c_m(\lambda_m - \lambda_{m-1})]^{-1}.$$

Proof. Let $\sum_{k=1}^m c_k \mathbf{W}_m^{(k)} = \mathbf{B} = (b_{ij})$.

We observe that in the matrix $\mathbf{B}$ $b_{1j} = b_{ij} = \sum_{k=j}^m c_k$ $(i = 2, 3, \cdots, m; j = i, i + 1, \cdots, m)$. Hence, according to a result given by Guttman [3], Ukita [4] (also stated in Greenberg and Sarhan [2]) $\mathbf{B}^{-1}$ exists and it is a diagonal matrix of type 2 (see [2] for definition). The elements of $\mathbf{B}^{-1} = (b^{ij})$ can be obtained by using a method given by Greenberg and Sarhan [2] as follows:

$$(5) \qquad \begin{aligned} &b^{11} = c_1^{-1}; \quad b^{12} = -c_1^{-1}; \quad b^{1j} = 0, \; j > 2. \\ &b^{i-1,i} = -c_{i-1}^{-1}; \; b^{ii} = c_{i-1}^{-1} + c_i^{-1}; \quad b^{i,i+1} = -c_i^{-1} \\ &b^{ij} = 0 \text{ for } j = 1, 2, \cdots, i-2, i+2, \cdots, m; i = 2, 3, \cdots, m. \end{aligned}$$

Since, $\mathbf{B}^{-1}$ is a diagonal matrix of type 2, $\mathbf{I}_m + \mathbf{B}^{-1}$ is also a matrix of the same

Received 16 November 1966; revised 3 March 1967.

[1] The research was sponsored by NSF Grant No. GK 695 at Columbia University.

type. Hence, it follows ([2], [3], [4]) that $(\mathbf{I}_m + \mathbf{B}^{-1})^{-1}$ exists and has the form given in equation (1). We also observe that

$$(6) \qquad (\mathbf{I}_m + \textstyle\sum_{k=1}^{m} c_k \mathbf{W}_m^{(k)})^{-1} = (\mathbf{I}_m + \mathbf{B})^{-1} = \mathbf{I}_m - (\mathbf{I}_m + \mathbf{B}^{-1})^{-1}.$$

We now form the matrix $\mathbf{I}_m + \mathbf{B}^{-1}$ from equation (5), multiply it by the matrix $\mathbf{Q}$ given in equation (1) and use equations (2), (3) and (4). This gives

$$(\mathbf{I}_m + \mathbf{B}^{-1})\mathbf{Q} = \mathbf{I}_m,$$

which completes the proof of the theorem.

It may be pointed out that the system of equations given in (2) can be solved recursively by first finding λ_3 from λ_2 and λ_1, then λ_4 from λ_3 and λ_2 and so on. We can then find the value of q_m from (4) explicitly. The system of equations given in (3) can then be solved successively by first finding q_{m-1} from q_m, then q_{m-2} from q_{m-1} and so on.

COROLLARY 1. *The quantities* λ_j, q_j $(j = 1, 2, \cdots, m)$ *satisfy the conditions*:

$$(7) \qquad \lambda_m > \lambda_{m-1} > \cdots > \lambda_2 > \lambda_1 = 1;$$

$$(8) \qquad q_j > 0 \qquad (j = 1, 2, \cdots, m).$$

PROOF. Let us assume that $\lambda_r > \lambda_{r-1} > 0$. Then, from (2)

$$(9) \qquad c_{r-1}(\lambda_{r+1} - \lambda_r) = c_r(\lambda_r - \lambda_{r-1}) + \lambda_r c_{r-1} c_r.$$

Equation (9) asserts that $\lambda_{r+1} > \lambda_r$ if $\lambda_r > \lambda_{r-1} > 0$. Since, $\lambda_2 > \lambda_1 > 0$, the first part of the corollary follows by induction. Also from (4) and (7), we obtain

$$(10) \qquad q_m > 0.$$

We now substitute $r = m$ in (3). This gives

$$(11) \qquad q_{m-1} = \lambda_m^{-1}\lambda_{m-1} q_m + c_{m-1}\lambda_m^{-1}.$$

Since $q_m > 0$, it follows from (11) that $q_{m-1} > 0$. We now substitute $r = m - 1, m - 2, \cdots, 2$ in (3) and use the above argument recursively to complete the proof of the second part of the corollary.

COROLLARY 2. *Suppose*, $c_1 = c_2 = \cdots = c_m = c$. *Then*, $\lambda_2, \lambda_3, \cdots, \lambda_m, q_1, q_2, \cdots, q_m$ *are given by*

$$\lambda_j = (1 + A^{2j-1})A^{1-j}(1 + A)^{-1} \qquad (j = 1, 2, \cdots, m)$$

where

$$A = 1 + c/2[1 + (1 + 4/c)^{\frac{1}{2}}],$$

$$q_m = A^m(1 + A)(1 + A^{2m+1})^{-1}c,$$

$$q_{m-1} = A^m(1 + A)(1 + A^{2m+1})^{-1}c(2 + c),$$

$$q_j = \lambda_j[q_{m-1}\lambda_{m-1}^{-1} - c(1 + A)(A - 1)^{-1}(1 + A^{2m-3})^{-1}] + c(A - 1)^{-1}A^{1-j} \qquad (j = 1, 2, \cdots, m - 2).$$

PROOF. When $c_1 = c_2 = \cdots = c_m = c$, equation (2) reduces to

$$\lambda_{r+1} - \lambda_r(2 + c) + \lambda_{r-1} = 0 \quad (r = 2, 3, \cdots, m - 1) \tag{12}$$

with $\lambda_1 = 1, \lambda_2 = 1 + c$.

The system of equations given in (12) can be solved by using a generating function and it is easy to see that

$$\lambda_j = [G^{(j)}(0)][j!]^{-1}, \tag{13}$$

where,

$$G(Z) = (Z - Z^2)[1 - (2 + c)Z + Z^2]^{-1}, \tag{14}$$

and $G^{(j)}(0)$ is the jth derivative of $G(Z)$ at $Z = 0$. Using the fact that A and A^{-1} are the roots of the quadratic $z^2 - (2 + c)z + 1$, we can write

$$\begin{aligned} G(Z) &= (Z - Z^2)[(Z - A)(z - A^{-1})]^{-1} \\ &= (Z - Z^2)[A(A^2 - 1)^{-1}(Z - A)^{-1} + A(1 - A^2)^{-1}(Z - A^{-1})^{-1}] \end{aligned}$$

whence,

$$\lambda_j = [G^{(j)}(0)][j!]^{-1} = (1 + A^{2j-1})A^{1-j}(1 + A)^{-1}. \tag{15}$$

This proves the first part of the corollary. Under the condition $c_1 = c_2 = \cdots = c_m = c$, we can rewrite equations (3) and (4) as

$$\lambda_{r-1}q_r - \lambda_r q_{r-1} + c = 0 \quad (r = 2, 3, \cdots, m), \tag{16}$$

$$q_m = c[\lambda_m(2 + c) - \lambda_{m-1}]^{-1}. \tag{17}$$

The expressions of λ_m and λ_{m-1} can be obtained from (15). This together with the condition $2 + c = A + A^{-1}$ lead, after some simple calculation, to

$$q_m = A^m(1 + A)(1 + A^{2m+1})^{-1}c. \tag{18}$$

The expression of q_{m-1} can be obtained by substituting $r = m$ in (16). Finally we substitute $r = m - 1, m - 2, \cdots, 3, 2$ successively in equation (16). After making some simple calculations, we obtain

$$q_j = \lambda_j\lambda_{m-1}^{-1}q_{m-1} + c\lambda_j \textstyle\sum_{k=j}^{m-2} \lambda_k^{-1}\lambda_{k+1}^{-1}. \tag{19}$$

Now,

$$\begin{aligned} \textstyle\sum_{k=j}^{m-2} \lambda_k^{-1}\lambda_{k+1}^{-1} &= (1 + A)^2 \textstyle\sum_{k=j}^{m-2} A^{2k-1}(1 + A^{2k-1})^{-1}(1 + A^{2k+1})^{-1} \\ &= (1 + A)(1 - A)^{-1} \textstyle\sum_{k=j}^{m-2} [(1 + A^{2k+1})^{-1} \\ &\quad - (1 + A^{2k-1})^{-1}] \\ &= (1 + A)(1 - A)^{-1}[(1 + A^{2m-3})^{-1} - (1 + A^{2j-1})^{-1}]. \end{aligned} \tag{20}$$

From (19) and (20)

$$q_j = \lambda_j[q_{m-1}\lambda_{m-1}^{-1} - c(1 + A)(A - 1)^{-1}(1 + A^{2m-3})^{-1}] + c(A - 1)^{-1}A^{1-j} \quad (j = 1, 2, \cdots, m - 2).$$

This completes the proof of the corollary.

We now consider an application of the result. Let $x_1, x_2, \cdots, x_n$ be a sequence of random variables such that

$$x_j = \mu_n + \epsilon_j + \textstyle\sum_{k=j}^{n-1} J_k z_k \qquad (j = 1, 2, \cdots, n-1),$$
$$= \mu_n + \epsilon_n \qquad (j = n),$$

with

$$E(\epsilon_k) = 0; \quad \text{Var}(\epsilon_k) = 1 \quad (k = 1, 2, \cdots, n),$$
$$E(z_k) = 0; \quad \text{Var}(z_k) = \sigma^2 \quad (k = 1, 2, \cdots, n-1),$$
$$P(J_k = 1) = P_k = 1 - P(J_k = 0) \quad (k = 1, 2, \cdots, n-1); 0 < P_k \leqq 1.$$

The variables ϵ_k $(k = 1, 2, \cdots, n)$, J_k, z_k $(k = 1, 2, \cdots, n-1)$ are assumed to be mutually independent. This model was introduced by Chernoff and Zacks [1] in connection with a problem when the means of a sequence of random variables are changing randomly. It is easy to see that the dispersion matrix $\mathbf{V}$ of $x_1, x_2, \cdots, x_n$ is

$$\mathbf{V} = \begin{pmatrix} \mathbf{I}_{n-1} + \sigma^2 \sum_{k=1}^{n-1} P_k \mathbf{W}_{n-1}^{(k)} & \mathbf{0} \\ \mathbf{0}' & 1 \end{pmatrix}$$

where $\mathbf{0}$ is an $n - 1 \times 1$ vector with all elements 0. The theorem discussed above can be applied with $m = n - 1$ and $c_k = \sigma^2 P_k$ $(k = 1, 2, \cdots, n-1)$ to obtain $\mathbf{V}^{-1}$. Chernoff and Zacks [1], p. 1003, gave an expression for the sum of the elements of each column of $\mathbf{V}^{-1}$ when P_k's are all equal. The present note generalizes this result firstly by deriving an expression of $\mathbf{V}^{-1}$ when P_k's are unequal and secondly by giving the complete matrix $\mathbf{V}^{-1}$ instead of the column sums only when P_k's are equal.

The author is grateful to the referees for several helpful remarks.

REFERENCES

[1] Chernoff, H. and Zacks, S. (1964). Estimating the current mean of a normal distribution which is subjected to changes in time. *Ann. Math. Statist.* **35** 999–1018.

[2] Greenberg, B. G. and Sarhan, A. E. Matrix inversion, its interest and application in analysis of data. *J. Amer. Statist. Assoc.* **54** 755–766.

[3] Guttman, L. (1955). A generalized simplex for factor analysis. *Psychometrika* **20** 173–195.

[4] Ukita, Y. (1955). Characterization of diagonal matrices with an application to order statistics. *J. Hokkaido College Art Literature* **6** 66–75.