A Note on Trace-Differentiation and the Ω-operator

By A. C. AITKEN

(*Received* 27*th January* 1949. *Read* 4*th March* 1949.)

1. In theory of polarizing operators in invariants the operator

$$\Omega \equiv \left[\frac{\partial}{\partial x_{ji}}\right], \tag{1}$$

where $X = [x_{ij}]$ is an $n \times n$ matrix of n^2 independent elements x_{ij}, holds an important place. Acting upon particular scalar functions of X, namely the spur or trace of powers of X, or of polynomials or rational functions of X with scalar coefficients, it exhibits (Turnbull, 1927, 1929, 1931) an exact analogy with results in the ordinary differentiation of the corresponding functions of one scalar variable. Turnbull denotes this operation of trace-differentiation under Ω by Ω_s; and we shall follow him. Our purpose is to show how, with a suitably modified Ω, the results may be extended to the case of symmetric matrices $X = X'$ having $\frac{1}{2}n(n+1)$ independent elements.

Such extensions forced themselves on the notice of the author in some work (Aitken, 1946) on the estimation of statistical parameters. At the same time and independently Gårding was using the determinant of the same modified Ω to obtain the analogue (Gårding, 1947) of Cayley's determinantal theorem (cf. Turnbull, 1928, p. 114)

$$|\,\Omega\,|\;|\,X\,|^r = r(r+1)(r+2)\ldots(r+n-1)\,|\,X\,|^{r-1}, \tag{2}$$

namely

$$|\,\Omega\,|\;|\,X\,|^r = r(r+\tfrac{1}{2})(r+1)\ldots(r+\tfrac{1}{2}n-\tfrac{1}{2})\,|\,X\,|^{r-1}, \tag{3}$$

where in (3) the matrix X is symmetric and Ω is modified.

2. We first establish the result 1 (2), following one of Turnbull's proofs, when r is a positive integer. The extension to the case of negative powers of X and to rational functions will then be taken in hand. The corresponding results for a symmetric matrix X follow without difficulty.

It is of advantage to use tensor notation and the summation convention. We denote x_{ij} by x^i_j, and we agree that the summation convention shall hold for repeated Greek indices, but not for italic.

https://doi.org/10.1017/S0013091500014152 Published online by Cambridge University Press

Thus the $(i, j)^{\text{th}}$ element in X^r, the i^{th} diagonal element, and the trace tr X^r, are respectively

$$x^i_\alpha x^\alpha_\beta x^\beta_\gamma \dots x^\pi_j, \qquad x^i_\alpha x^\alpha_\beta x^\beta_\gamma \dots x^\pi_i, \qquad \text{tr}\, X^r = x^\rho_\alpha x^\alpha_\beta x^\beta_\gamma \dots x^\pi_\rho, \tag{1}$$

the last being an aggregate of n^r terms, the row and column indices of the factors in each term being linked, as shown, in a closed cycle of r indices. The sub-aggregate of terms in tr X^r having x^j_i for first factor is $x^j_i x^i_\beta x^\beta_\gamma \dots x^\pi_j$.

To differentiate partially with respect to x^j_i is to delete the factor x^j_i once wherever it occurs, and to sum all terms so obtained, each such term being a product of $r - 1$ elements. Performing this deletion of x^j_i in each of the r factors we obtain r times in all, in virtue of the cyclic symmetry, the expression $x^i_\beta x^\beta_\gamma \dots x^\pi_j$, that is, the $(i, j)^{\text{th}}$ element in X^{r-1}. We thus have, for this first case of a general matrix X,

$$\Omega_s X^r = r X^{r-1}, \tag{2}$$

and the extension

$$\Omega_s f(X) = f'(X), \tag{3}$$

where $f(X)$ is a polynomial in X with scalar coefficients, is immediate.

Further extensions, such as

$$\Omega_s A X^r = A X^{r-1} + X A X^{r-2} + X^2 A X^{r-3} + \dots + X^{r-1} A, \tag{4}$$

admit of similar proof. We have only to notice that the trace of AX^r is $a^\sigma_\alpha x^\alpha_\beta x^\beta_\gamma \dots x^\rho_\sigma$, to carry out deletion of x^j_i in each of the r positions and to observe the fact of cyclic symmetry. More generally, and for similar reasons, the result of Ω_s acting on $AX^pBX^qCX^r \dots X^tK$ is given by the rule: Arrange the factors of the operand, X^p being written out as $XXX \dots X$ and so on, in all the possible cyclic orders in which X appears as first factor. In each of these, delete this first X. The sum of the resulting products of matrices is the desired trace-derivative.

3. We now extend the result 2 (2) to the case X^{-r}, where r is a positive integer. The $(i, j)^{\text{th}}$ element in X^{-r} is then $X^j_\alpha X^\alpha_\beta X^\beta_\gamma \dots X^\pi_i$ and so

$$\text{tr}\, X^{-r} = X^\rho_\alpha X^\alpha_\beta X^\beta_\gamma \dots X^\pi_\rho. \tag{1}$$

The effect of $\dfrac{\partial}{\partial x_{ji}}$ is, as usual, to operate on each factor as if the rest were constant, and to add the results. Denoting the cofactor of x^j_i in $| X^\alpha_\beta |$ by $| X^{\alpha j}_{\beta i} |$, and observing that $X^\alpha_\beta = | X^\alpha_\beta | / | X |$,

https://doi.org/10.1017/S0013091500014152 Published online by Cambridge University Press

we have

$$\frac{\partial}{\partial x_{ji}} X^{\alpha}_{\beta} = \begin{vmatrix} X^{\alpha j}_{\beta i} & X^{\alpha}_{\beta} \\ X^{j}_{i} & X \end{vmatrix} \div | X |^{2}. \tag{2}$$

Now the numerator is an "extensional" determinant, and so (2) is readily evaluated as

$$- X^{\alpha}_{i} \ldots X^{j}_{\beta}. \tag{3}$$

Thus whereas in case of X^r trace-differentiation has the effect of *deleting* in r ways a link from a chain, in the case of X^{-r} it has the effect of *inserting* a link, again in r ways, with change of sign. For example one of the r terms in

$$\frac{\partial}{\partial x_{ji}} X^{\rho}_{\alpha} X^{\alpha}_{\beta} X^{\beta}_{\gamma} \ldots X^{\pi}_{\rho} \qquad \text{is} \qquad - X^{\rho}_{\alpha} X^{\alpha}_{i} X^{j}_{\beta} X^{\beta}_{\gamma} \ldots X^{\pi}_{\rho},$$

namely a term in the $(i, j)^{\text{th}}$ element of $- X^{-r-1}$. Thus, again in virtue of cyclic symmetry, we have

$$\Omega_s X^{-r} = - r X^{-r-1}. \tag{4}$$

It is easy to deduce a result analogous to 2 (4) for the effect of Ω_s on AX^{-r}.

4. Next let X be a *symmetric* matrix $X = X'$, the elements being $\frac{1}{2}n(n+1)$ independent variables. The main lines of the reasoning are the same; we have only to note the alteration caused by the fact that $x_{ji} = x_{ij}$. First, deletions and insertions with respect to x^{i}_{i} will be as before. With respect to x^{j}_{i}, $i \neq j$, to every term obtained by a deletion (or insertion) with respect to it there will correspond, by symmetry, a duplicate term obtained by a deletion (or insertion) with respect to x^{i}_{j}. Thus the sets of terms arising from the non-diagonal operators in Ω_s will appear twice. We can compensate for this by using a suitably modified Ω, namely

$$\Omega = \left[\epsilon_{ji} \frac{\partial}{\partial x_{ji}} \right], \qquad \epsilon_{ii} = 1, \ \epsilon_{ij} = \epsilon_{ji} = \tfrac{1}{2}, \ i \neq j, \tag{1}$$

and now we have, for a symmetric matrix X,

$$\Omega_s X^{r} = r X^{r-1}, \qquad \Omega_s X^{-r} = - r X^{-r-1}. \tag{2}$$

We may add, for completeness, the result

$$\Omega_s \log X = \Omega \mid X \mid = X^{-1}, \qquad | X | \neq 0, \tag{3}$$

easily proved in an elementary way.

https://doi.org/10.1017/S0013091500014152 Published online by Cambridge University Press

The case of a rational function $f(X)\{g(X)\}^{-1}$ can be treated by resolution into partial fractions and use of the result, an immediate consequence of 2 (2),

$$\Omega_s (X + aI)^{-r} = - r (X + aI)^{-r-1}, \tag{4}$$

where a is a scalar constant.

Thus, both in the general and in the symmetric case, we have theorems, in trace-differentiation of polynomials and rational functions of X, resembling in all respects the familiar theorems of ordinary differentiation.

5. It is of interest, even though the results are rather formal, to note that with a further slight modification of Ω we may apply trace-differentiation to skew symmetric matrices. We then have

$$X' = - X, \text{ that is, } x_{ii} = 0,\ x_{ji} = - x_{ij},$$

and it is readily shown that if

$$\Omega = \left[\epsilon_{ji} \frac{\partial}{\partial x_{ji}}\right],\ \epsilon_{ii} = 0,\ \epsilon_{ij} = \epsilon_{ji} = \tfrac{1}{2},\ i \neq j,$$

then trace-differentiation preserves those formal analogies which it has been the object of this note to communicate. It must be observed, however, that odd powers of X are debarred, since the trace is then identically zero. The rule applies to rational functions in which only even powers appear.

REFERENCES.

Aitken, A. C., *Proc. Roy. Soc. Edinburgh*, 62 (1948), 369-377, 374.

Gårding, L., *Proc. Edinburgh Math. Soc.* (2), 8 (1948), 73-75.

Turnbull, H. W., *Proc. Edinburgh Math. Soc.* (2), 1 (1927), 111-128 ; (2), 2 (1930), 33-54,[1] 256-264 ; (2), 8 (1948), 76-86 ; *Theory of Determinants, Matrices and Invariants* (1928), 72, 114-115, 123.

[1] In (5) and (6) on p. 42 of this paper, $- rX^{-r+1}$ should be replaced by $- rX^{-(r+1)}$.

MATHEMATICAL INSTITUTE,
UNIVERSITY OF EDINBURGH.

https://doi.org/10.1017/S0013091500014152 Published online by Cambridge University Press