PUBLICATIONS DE L'INSTITUT MATHÉMATIQUE
Nouvelle série, tome 21 (35), 1977, pp. 165–168

ON DEFINITENESS OF CERTAIN MATRIX FUNCTIONS OF A MATRIX

Bojan Popović

(Received November 8, 1976)

0. The best known matrix functions [3, Chs. 6 and 11] have square matrices of the same dimensions (generalizing some real functions) or scalars (e.g. the coefficients of the characteristic polynomial) for their values. Other types of matrix functions have been formulated only for some special purposes, e.g. [1]. In this paper another class of matrix functions, appearing in some vector optimal adaptive control and nonlinear filtering problems [6], is defined as a particular mapping of square matrices into square matrices of smaller dimensions in general. Some properties of these functions are shown and two sufficient conditions for the definiteness of their image are formulated.

1. Real matrix M of the dimensions (m_1, m_2); $m_1, m_2 \geqslant 1$ is an $(m_1 \times m_2)$ ordered array $(m_{ij})_j^i$ of the real elements m_{ij}; $i=1, \ldots, m_1$, $j=1, \ldots, m_2$. For $m_2=1$ it is called m_1-vector and, for emphasis, denoted $\boldsymbol{m}$; for $m_1=m_2=1$ it is taken isomorphic to the corresponding scalar m_{11}. For the square matrix $M\,(m_1=m_2 \stackrel{*}{=} m)$ of order $\operatorname{ord} M = m$ its trace is a scalar function with the value

$$\operatorname{tr} M \stackrel{*}{=} \sum_{i=1}^{\operatorname{ord} M} m_{ii}. \tag{1}$$

D e f i n i t i o n. *For each positive integer n a function T_n is defined for any square matrix $M \stackrel{*}{=} (m_{ij})_j^i$ of order m that is a multiple of n, i.e. $m = n \cdot r$ for some positive integer r, and its value is a square matrix of order r*

$$T_n M \stackrel{*}{=} \Big(\sum_{k=1}^{n} m_{(i-1)n+k,\,(j-1)n+k}\Big)_j^i. \tag{2}$$

In view of (1)

$$T_n M = (\operatorname{tr} M_{ij})_j^i. \tag{2'}$$

where M_{ij}; $i, j = 1, \ldots, r$ is a square submatrix of (partitioned) matrix M such that its elements are $m_{(i-1)n+k,\,(j-1)n+l}$; $k, l = 1, \ldots, n$. E.g.

$$T_2 \left[\begin{array}{cc|cc} 1 & 1 & 2 & 0 \\ 3 & 0 & 0 & 0 \\ \hline 2 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] = T_2 \left[\begin{array}{cc|cc} 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 2 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] = T_2 \left[\begin{array}{cc|cc} -1 & 0 & 2 & 0 \\ 0 & 2 & 0 & 0 \\ \hline 2 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}.$$

Generally, the function T_n is shrinking a square matrix into a smaller one by replacing every submatrix of order n in a symmetric partitioning by its scalar trace value.

2. The functions T_n are, in a sense, generalizations of a trace since this function is used in forming their value, cf. (2′), and, furthermore, $T_{\operatorname{ord} M} M \equiv \operatorname{tr} M$. At the other extreme value of the index set $\{n \mid \operatorname{ord} M/n \text{ is a positive integer}\} \subseteq \{1, \ldots, \operatorname{ord} M\}$ $T_1 M \equiv M$.

For $n > 1$ functions T_n are many-to-one mappings since then there are infinitely many matrices M_{ij} of order n with the same trace, cf. the example in 1.

These functions have the following properties:

$$T_n M' = (T_n M)', \tag{3}$$

$$T_n (N + cM) = T_n N + c T_n M \tag{4}$$

if the left-hand side exists for scalar $c \neq 0$, both by elementary properties of the trace; and

$$T_{n \cdot n'} M = T_n (T_{n'} M) \tag{5}$$

if the left-hand side exists for positive integers n and n', by (2) since $\sum_{l=1}^{n} \sum_{k=1}^{n'} (k-1)\, n + l = \sum_{k=1}^{n \cdot n'} k$.

3. From (3) it is implied that $T_n M$ is symmetric for symmetric matrices M,[1] therefore the following theorem is well-posed.

Theorem A. *If M is a non-negative (positive) definite matrix of order m, then $T_n M$ of order $r \doteq m/n$ for each n such that $m \pmod n = 0$, is also a non--negative (positive) definite matrix.*[1]

Proof.[2] For $n = 1$ the theorem is trivial. For $n > 1$ the proof is based on observation that the sum of n conveniently chosen square submatrices around the main diagonal (principal minors) of order $m - n + 1$ of matrix M has among its elements all the elements of matrix $T_n M$. Since M is a non-negative (positive) definite matrix if and only if the determinants of all[3] its principal minors, $M(p_1, \ldots, p_k) \doteq (m_{p_i p_j})^i_j$; $1 \leqslant p_1 < \cdots < p_k \leqslant m$, $k = 1, \ldots, m$, are non-negative (positive) [4, p.307], then the principal minors of M are of the same definiteness as M. In particular, matrices $M(k, k+1, \ldots, k+m-n)$; $k = 1, \ldots, n\,(>1)$ of order $m - n + 1 = (r-1)\,n + 1$ are non-negative (positive) definite together with M, and so is their sum, $T \doteq \sum_{k=1}^{n} M(k, k+1, \ldots, k+m-n)$. Hence, the following principal minor of order r of the matrix T, $T(1, 1+n, \ldots, 1+(r-1)\,n) = \left(\sum_{k=1}^{n} m_{k-1+(1+(i-1)n),\, k-1+(1+(j-1)n)}\right)^i_j$, which is equal to $T_n M$ by (2), is also a non-negative (positive) definite matrix. Q.E.D.

---

[1] Clearly, for $n > 1$ the converse does not hold for every matrix M, cf. the example in 1.

[2] In view of (5) it would suffice to prove it only for prime numbers n.

[3] For positive definiteness only $p_l = l$; $l = 1, \ldots, k$ is sufficient [3, p.74].

Since by (4) $T_n(-M) = -T_n M$, Theorem A holds if the words "negative" and "positive" are interchanged.

4. If M and N are two non-negative definite matrices of the same order m, then $T_n NM$ is not necessarily a non-negative definite matrix in general, e.g. for $n=1$, except for $n=m$ $(=\operatorname{ord} NM)$: $T_m NM \equiv \operatorname{tr} NM = \sum_{i,j=1}^{m} n_{ij}\cdot m_{ji} \geqslant 0$ [3, p.102]. But if N is of special quasi-scalar form $I_r \otimes P = \begin{bmatrix} P & & 0 \\ & \ddots & \\ 0 & & P \end{bmatrix}$,[4] where P is a non-negative definite matrix of order n and the positive integer $r = m/n$, the above statement is true for all n, i.e.

Theorem B. *If P and M are both non-negative (non-positive) definite matrices of orders n and m, respectively, where $m \pmod n = 0$, then $T_n(I_r \otimes P)M$ is a non-negative definite matrix when I_r is the identity matrix of order $r \mathrel{\underline{*}} m/n$ and $\otimes$ denotes the (right direct) Kronecker product of matrices.*

Proof. If P is a non-negative definite matrix, a square matrix Q of the same order exists such that $P = Q'Q$ [3, p.54]. Hence, if M is a non-negative definite matrix,

$$T_n(I_r \otimes P)M = (\operatorname{tr} PM_{ij})_j^i = (\operatorname{tr} Q' M_{ij} Q)_j^i = T_n(I_r \otimes Q)' M (I_r \otimes Q)$$

is also a non-negative definite matrix by Theorem A, since the matrix $A'MA$ is then non-negative definite for any matrix A [4, p.305] and $\operatorname{tr} AB = \operatorname{tr} BA$ whenever both sides coexist [3, p.95].

Since $I_r \otimes (-P))(-M) \equiv (I_r \otimes P)M$, the theorem holds also for the non--positive definite matrices P and M. Q.E.D.

5. The functions T_n were initiated by the following identity [6, p.152] for r-vector $\mathbf{u}$

$$\operatorname{tr}(\boldsymbol{u}' \otimes I_n) M (\boldsymbol{u} \otimes I_n) P \equiv \boldsymbol{u}'(T_n(I_r \otimes P)M)\boldsymbol{u}, \tag{6}$$

This quadratic form appears in the Hamiltonian for derivation of r-vector $\boldsymbol{u}$ optimal control of the stochastic linear dynamics with n-vector state $\boldsymbol{x}$ and the "random gain" matrix $B \mathrel{\underline{*}} (b_{ij})_j^i$ of the dimensions (n, r):

$$\boldsymbol{x}(t+1) = A\boldsymbol{x}(t) + B\boldsymbol{u}(t) + \boldsymbol{e}(t) \equiv A\boldsymbol{x}(t) + (\boldsymbol{u}'(t) \otimes I_n)\boldsymbol{z} + \boldsymbol{e}(t),$$

where $\boldsymbol{z}' \mathrel{\underline{*}} (b_{11}, \ldots, b_{n1}, b_{12}, \ldots, b_{nr})$, i.e. $\boldsymbol{z}$ is $n\cdot r$-vector — a column by column of elements of the random matrix B — so that its second-order central moment is representable by the covariance matrix $C(\boldsymbol{z})$ of order $n\cdot r \mathrel{\underline{*}} m$ [5]. If the covariance matrix of $\boldsymbol{x}(t)$ is denoted by $C(\boldsymbol{x}, t)$ — the matrix of order n, then the Hamiltonian for the adaptive optimal vector controls (generalizing [7]) contains [6, p.80] the quadratic form (6) for non-negative definite matrices $M \equiv C(\boldsymbol{z})$ and $P \equiv P(C(\boldsymbol{x}, t+1))$, costate of $C(\boldsymbol{x}, t+1)$.[5] Hence, the sufficient condition for optimum control $\boldsymbol{u} \equiv \boldsymbol{u}(t)$ could be satisfied by Theorem B, e.g. in the

[4] (Right direct) Kronecker product is defined by $A \otimes B \mathrel{\underline{*}} (a_{ij}B)_j^i$ [3, p.227].

[5] $C(\boldsymbol{z})$ is alaways non-negative definite as a covariance matrix, and this holds also for $P(C(\boldsymbol{x}, t+1))$ if the final boundary condition on $P(C(\boldsymbol{x}, .))$ is non-negative definite and that is fulfilled for any reasonable cost functional.

usual case when the time-additive cost functional for the problem contains positive definite quadratic form in controls, $\boldsymbol{u}'(t)\, V(t)\, \boldsymbol{u}(t)$.[6]

A similar form and the condition [6, pp. 173—175] appears in deriving the (near) optimal (in a sense of minimal covariance unbiased, i.e. zero-mean error) filter (generalizing [2]) for nonlinear state dynamics controlled systems [6, Appendix E].

6. The author is indebted to Professor Franz E. Hohn of the University of Illinois for the initial encouragement to write a paper and to Professor Slaviša B. Prešić of the University of Beograd for helpful suggestions on an earlier draft.

---

[6] The second derivative of the Hamiltonian with respect to $u(t)$ is then a positive definite matrix $V(t)+T_n(I_r \otimes P(C(x, t+1)))\, C(z)$.

REFERENCES

[1] Afriat, S. N., *Composite matrices*, Quartely Journal of Mathematics, 5 (second series), 81—98 (1954).

[2] Athans, M., R. P. Wisher, and A. Bertolini, *Suboptimal state estimation for continuous-time nonlinear systems from discrete noisy measurements*, IEEE Transactions on Automatic Control, AC—13 (5), 504—514 (1968).

[3] Bellman, R., *Introduction to matrix analysis*. New York: Mc Graw-Hill, 1960.

[4] Gantmacher, F. R., *The theory of matrices*. vol I, New York, Chelsea Publ. Co., 1959.

[5] Nissen, D. H., *A note on the variance of a matrix*, Econometrica, 36 (3—4) 603—604 (1968).

[6] Popović, B., *Analytic solutions to some stochastic, adaptive, and dual optimal control problems for economic decision making*, Urbana, Ill., CSL Report T-4, 1972.

[7] Tse, E. and M. Athans, *Adaptive stochastic control for a class of linear systems*, IEEE Transactions on Automatic Control, AC—17 (1), 38—51 (1972).

Address of the author:
Ekonomski institut
M. Tita 16
11 000 Beograd, Yugoslavia