Characterizations of Products of Symmetric Matrices

Chi Song Wong
University of Windsor
Windsor, Ontario N9B 3P4, Canada

Submitted by Ingram Olkin

ABSTRACT

Characterizations are obtained for matrices C of the form $C=A\Sigma$, where A, Σ are $n\times n$ matrices over the real field such that A is symmetric and C is nonnegative definite. Among others, a proof of recent generalization of Cochran's theorem is given.

1. INTRODUCTION

We shall be interested only in matrices over the real field R. $M_{m\times n}$ will be the set of $m\times n$ matrices. W_n will be the set of all $C=A\Sigma$, where $A, \Sigma\in M_{n\times n}$, A is symmetric, and Σ is nonnegative definite. Among other things, we shall characterize W_n and give a different proof of a generalization of Cochran's theorem [2]. Matrices in W_n occur in linear models and multivariate analysis, where Σ is the dispersion matrix of a normal random vector X, and A is determined by a given quadratic form $Y=X'AX$ of X [1, 9].

2. CHARACTERIZATIONS OF CERTAIN CLASSES OF MATRICES

Since A' is similar to A, $\begin{pmatrix} D & 0 \\ E & 0 \end{pmatrix}$ below can be replaced by $\begin{pmatrix} D & E' \\ 0 & 0 \end{pmatrix}$.

PROPOSITION 2.1. *Let $C\in M_{n\times n}$. Then $C=A\Sigma$ for some symmetric matrix A in $M_{n\times n}$ and some nonnegative definite matrix Σ of rank s if and only if C is similar to a matrix of the form*

$$\begin{pmatrix} D & 0 \\ E & 0 \end{pmatrix},$$

where D is a diagonal matrix in $M_{s\times s}$ and $E\in M_{(n-s)\times s}$.

52 Vanderbilt Ave., New York, NY 10017 0024-3795/82/010243+9$02.75

Proof. *Only if*: By changing A to $P'AP$ for an orthogonal P such that $P\Sigma P'$ is diagonal, we may assume that $\Sigma=\text{diag}(G_s,0)$, where $G_s\in M_{s\times s}$ is positive definite. Write

$$A=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix},$$

where A_{11} is an $s\times s$ matrix. Then A_{11} is symmetric and

$$C=\begin{pmatrix} A_{11}G_s & 0 \\ A_{21}G_s & 0 \end{pmatrix}.$$

Let $K=\text{diag}(G_s^{1/2}, I_{n-s})$. Then

$$KCK^{-1}=\begin{pmatrix} G_s^{1/2}A_{11}G_s^{1/2} & 0 \\ A_{21}G_s^{1/2} & 0 \end{pmatrix}.$$

Since $G_s^{1/2}A_{11}G_s^{1/2}$ is symmetric, there exist an $s\times s$ orthogonal matrix Q and a diagonal matrix D in $M_{s\times s}$ such that $G_s^{1/2}A_{11}G_s^{1/2}=QDQ'$. Let $W=\text{diag}(Q, I_{n-s})$. Then KCK^{-1} is similar to

$$W^{-1}KCK^{-1}W=\begin{pmatrix} D & 0 \\ E & 0 \end{pmatrix},$$

where $E=A_{21}G_s^{1/2}Q$. Thus C is similar to

$$\begin{pmatrix} D & 0 \\ E & 0 \end{pmatrix}.$$

If: By hypothesis,

$$C=P\begin{pmatrix} D & 0 \\ E & 0 \end{pmatrix}P^{-1}$$

for some nonsingular matrix P in $M_{n\times n}$. Since D is diagonal, $D=(d_i\delta_{ij})$ for some real numbers d_i, where δ_{ij} is the Kronecker symbol. Let

$$F=(f_i\delta_{ij})\in M_{s\times s}, \qquad G=(g_i\delta_{ij})\in M_{s\times s},$$

where $f_i = 1$, $g_i = d_i$ when $d_i > 0$, and $f_i = 0$, $g_i = 1$ when $d_i = 0$. Then

$$\begin{pmatrix} D & 0 \\ E & 0 \end{pmatrix} = \begin{pmatrix} F & G^{-1}E' \\ EG^{-1} & 0 \end{pmatrix} \begin{pmatrix} G & 0 \\ 0 & 0 \end{pmatrix}.$$

Let

$$A = P \begin{pmatrix} F & G^{-1}E' \\ EG^{-1} & 0 \end{pmatrix} P', \qquad \Sigma = (P')^{-1} \begin{pmatrix} G & 0 \\ 0 & 0 \end{pmatrix} P^{-1}.$$

Then $C = A\Sigma$, A is symmetric, and Σ is nonnegative definite of rank s. ■

COROLLARY 2.2. *Let $C \in M_{n \times n}$. Then $C = A\Sigma$ for some symmetric matrix A and positive definite matrix Σ if and only if C is similar to a diagonal matrix D.*

The following result follows from the above proof of Proposition 2.1.

PROPOSITION 2.3. *Let $C \in M_{n \times n}$. Then $C = A\Sigma$ for some nonnegative (positive) definite matrix A and positive definite matrix Σ if and only if C is similar to a nonnegative (positive) definite diagonal matrix D.*

We note here that even for $n = 2$, W_2 contains many matrices that are not similar to a diagonal matrix. For example, let

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

$C = A\Sigma$. Then by definition, $C \in W_2$. By a direct calculation,

$$C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

(and therefore is not symmetric). Since $C \neq 0$ and $C^2 = 0$, C is not similar to a diagonal matrix.

We now give a characterization of W_n in terms of Jordan forms.

PROPOSITION 2.4. *Let $C \in M_{n \times n}$. Then $C \in W_n$ if and only if each Jordan block (in the Jordan form) of C is either a real number or the two by two matrix*

$$Q = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Proof. *If*: Let $J=\text{diag}(J_1, J_2, \ldots, J_s)$ be the Jordan form of C, where the J_i's are all the Jordan blocks of C. Then $C=PJP^{-1}$ for some nonsingular matrix P. We may assume that $J_1, J_2, \ldots, J_t$ are real numbers and $J_{t+1}, J_{t+2}, \ldots, J_s$ are equal to Q. Now let

$$A_i=J_i, \quad \Sigma_i=1, \qquad i=1,2,\ldots,t,$$

$$A_i=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \Sigma_i=\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad i=t+1, t+2, \ldots, s,$$

$$A=P\text{diag}(A_1, A_2, \ldots, A_s)P', \qquad \Sigma=(P')^{-1}\text{diag}(\Sigma_1, \Sigma_2, \ldots, \Sigma_s)P^{-1}.$$

Then A is symmetric, Σ is nonnegative definite, and $C=A\Sigma$, i.e. $C\in W_n$.

Only if: By Proposition 2.1, there exists a nonsingular P in $M_{n\times n}$ such that

$$C=P\begin{pmatrix} D & 0 \\ E & 0 \end{pmatrix}P^{-1}$$

for some diagonal D in $M_{s\times s}$ and E in $M_{(n-s)\times s}$. Let

$$B_1=\begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix}, \quad B_2=\begin{pmatrix} 0 & 0 \\ E & 0 \end{pmatrix}, \qquad B=B_1+B_2.$$

Then $C=PBP^{-1}$, $B_1B_2=0$, and $B_2^2=0$. So $B^j=BB_1^{j-1}$, $j=1,2,\ldots$. Let

$$\psi_1(\lambda)=\sum_{i=0}^{p} a_i\lambda^i \qquad (a_p=1)$$

be the minimal polynomial of B_1. Since B_1 is diagonal,

$$\psi_1(\lambda)=\prod_{i=1}^{p}(\lambda-\lambda_i),$$

where the λ_i's are the distinct diagonal entries in B_1. In particular, λ^2 does not divide $\psi_1(\lambda)$. Let $\psi_2(\lambda)=\lambda\psi_1(\lambda)$. Then λ^3 does not divide $\psi_2(\lambda)$, and

$$\psi_2(B)=\sum_{i=0}^{p} a_iB^{i+1}=\sum_{i=0}^{p} a_iBB_1^i=B\psi_1(B_1)=0.$$

So the minimal polynomial $\psi(\lambda)$ of B divides $\psi_2(\lambda)$. Thus:

(a) Any Jordan block J_i of B corresponding to a nonzero λ_i has to be the real number λ_i, a 1×1 matrix.

(b) λ^3 does not divide $\psi(\lambda)$. So if $\lambda_0=0$ is an eigenvalue of B, then each Jordan block J_i of B corresponding to λ_0 must be either the 1×1 matrix (0) or the 2×2 matrix Q [5].

Since C is similar to B, (a) and (b) above complete the proof. ∎

The above result provides some information about W_n:

(1) If $C\in W_n$ and if J is the Jordan form of C, then the rank of C is equal to the number of nonzero entries in J.

(2) If $C\in W_n$ and if j is even, then C^j is similar to a diagonal matrix.

(3) $A\in W_n$ if and only if $A'\in W_n$.

(4) The family of nondiagonalizable matrices in W_n can be described. For example, $C\in W_2$ is not diagonalizable if and only if C is similar to Q.

(5) Even for $n=2$, $M_{n\times n}$ contains a lot of matrices which are not in W_n. Indeed $C\in M_{2\times 2}\backslash W_2$ and C has a Jordan form in $M_{2\times 2}$ if and only if C is similar to $\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix}$ for some nonzero a.

The following fairly strong decomposition about $C\in W_n$ also follows easily from the above proposition.

PROPOSITION 2.5. *Let $C\in M_{n\times n}$. Then $C\in W_n$ if and only if there exist $A, B\in M_{n\times n}$ such that*

(a) *there exists a nonsingular P in $M_{n\times n}$ such that $P^{-1}AP$, $P^{-1}BP$ are the Jordan forms of A, B respectively,*

(b) *A similar to a diagonal matrix,*

(c) *$B^2=0$,*

(d) *$AB=0$,*

(e) *$C=A+B$.*

The above result tells us that $C\in W_n$ is at most, a little bit $(=B)$ short of being diagonalizable. The following also follows easily from Proposition 2.4.

PROPOSITION 2.6. *Let $C\in M_{n\times n}$. Then*

(a) *$C\in W_n$ if $C^3=C^2$,*

(b) *$C^3=C^2$ if $C\in W_n$ and $\sigma(C)\subset\{0,1\}$,*

(c) *$C^4=C^2$ if $C\in W_n$ and $\sigma(C)\subset\{0,1,-1\}$.*

Idempotents arise in the study of linear models and Cochran's theorem [3] (e.g., see [6, 7, 10, 11, 2]). The following known result shows that all idempotents are in W_n. For any linear function f on a linear space R^n, denote by F_f the set $\{x: f(x)=x\}$ of all fixed points of f, by $\ker f$ the kernel $\{x: f(x)=0\}$ of f, and by $\operatorname{Im} f$ the image set $\{f(x): x\in R^n\}$ of f. Let $C\in M_{n\times n}$. We shall treat C as the linear transformation f: $f(x)=Cx$, $x\in R^n$. $\operatorname{Im} C$ is easily seen to be the column space of C. The dimension of a linear space L is denoted by $\dim L$.

PROPOSITION 2.7. *Let $C\in M_{n\times n}$. Then the following conditions are equivalent*:

(a) $C^2=C$.
(b) $R^n=F_c+\ker C$.
(c) *C is similar to a diagonal idempotent D.*
(d) $F_C=\operatorname{Im} C$.

It is hinted here idempotence is related to fixed points. This observation can be expanded. If C is an idempotent, then C is a generalized inverse C^- of C. Let $H\in M_{m\times m}$, $K\in M_{n\times m}$. Then H is a generalized inverse of K if and only if $F_{KH}=\operatorname{Im} K$. In other words, how far a generalized inverse H of K is from being an inverse of K can be measured by the size of F_{KH}. In particular, if $m=n$ and K has an inverse H, then $F_{KH}=R^n$, which is of the largest possible size. On the other hand, if $K=0$, then any $H\in M_{m\times n}$ is a generalized inverse of K and $F_{KH}=\{0\}$, which is of the smallest possible size and suggests that H may not be much of an inverse of K.

To see how C in Proposition 2.3 occurs in statistics, we prove the following. We emphasize here that unlike Searle [9, p. 57], Good [4], or Nagase and Banerjee [8], we do not assume that $X'AX$ has a central chi squared distribution.

PROPOSITION 2.8. *Let $A,\Sigma\in M_{n\times n}$ such that A is symmetric and Σ is nonnegative definite. Let X be a normal random vector with parameter* $(0,\Sigma)$ *such that $Y=X'AX$ has a chi squared distribution. Then $\sigma(A\Sigma)\subset\{0,1\}$.*

Proof. Since Y has a chi squared distribution, for some $m\in Z^+$, $s\geq 0$, the rth cumulant K_r of Y is [9]

$$K_r=2^{r-1}(r-1)!(m+2sr),\qquad r=1,2,\dots .$$

Since X is normal, $K_r=2^{r-1}(r-1)!\operatorname{tr}((A\Sigma)^r)$ [9], which implies that

$$m+2sr=\operatorname{tr}\big((A\Sigma)^r\big)=\sum_{i=1}^{n}\lambda_i^r, \tag{2.1}$$

where $\{\lambda_i\}$ is the spectrum of $A\Sigma$. If one of $|\lambda_i|$, say $|\lambda_{i_0}|$, is greater than 1, then by (2.1),

$$s \geqslant \lim_{p\to\infty} \frac{\lambda_{i_0}^{2p} - m}{4p} = \infty,$$

a contradiction. So all $\lambda_i \in [-1,1]$. By (2.1),

$$s = \lim_{r\to\infty} \frac{\sum_{i=1}^{n} \lambda_i^r - m}{2r} = 0,$$

so that $\sigma(A\Sigma)\subset\{0,1\}$. ■

Good concluded in [4] that $A\Sigma$ above is an idempotent. Styan [10] pointed out that $A\Sigma$ need not be idempotent. Now by Proposition 2.4, we can give a characterization of $A\Sigma$ under which $A\Sigma$ is an idempotent.

PROPOSITION 2.9. *Let $C\in W_n$ such that $\sigma(C)\subset\{0,1\}$. Then the following conditions are equivalent*:

(a) *C is an idempotent.*
(b) *C is diagonalizable.*
(c) *$r(C)=\operatorname{tr}C$.*
(d) *$r(C)=r(C^2)$.*

We now investigate conditions on A and/or Σ under which $A\Sigma$ is an idempotent. In this regard, we obtain the following result. Recall that for any $X\in M_{n\times m}$, $\ker X'X=\ker X$, so that we can cancel X' whenever $X'XA=X'XB$.

PROPOSITION 2.10. *Let $A,\Sigma\in M_{n\times n}$ such that A and Σ are nonnegative definite and $\sigma(A\Sigma)\subset\{0,1\}$. Then $A\Sigma$ is an idempotent.*

Proof. By Proposition 2.6, $A\Sigma A\Sigma A\Sigma = A\Sigma A\Sigma$. By canceling $A\Sigma^{1/2}$ we obtain $\Sigma^{1/2}A\Sigma A\Sigma=\Sigma^{1/2}A\Sigma$. Again, canceling $\Sigma^{1/2}A^{1/2}$, we obtain $A^{1/2}\Sigma A\Sigma = A^{1/2}\Sigma$. Multiplication by $A^{1/2}$ on the left yields $A\Sigma A\Sigma = A\Sigma$. ■

If one is interested in C in W_n with $C^4=C^2$ (see e.g. Tan [11]) then C^2 is diagonalizable. In fact C^2 is diagonalizable for all $C\in W_n$.

PROPOSITION 2.11. *Let $A,B,C\in M_{n\times n}$ such that $A=B+C$, $A^2=A$, and $r(A)=r(B)+r(C)$. Then B,C are idempotents and $BC=CB=0$.*

Proof.

$$r(A)=\dim \operatorname{Im} A \leqslant \dim(\operatorname{Im} B+\operatorname{Im} C)$$
$$=\dim \operatorname{Im} B+\dim \operatorname{Im} C-\dim \operatorname{Im} B \cap \operatorname{Im} C$$
$$=r(B)+r(C)-\dim \operatorname{Im} B \cap \operatorname{Im} C.$$

Since $r(A)=r(B)+r(C)$,

$$\operatorname{Im} B \cap \operatorname{Im} C=\{0\} \tag{2.2}$$

and

$$\operatorname{Im} A=\operatorname{Im} B+\operatorname{Im} C. \tag{2.3}$$

Let $x \in \operatorname{Im} B$. By (2.3), $x \in \operatorname{Im} A$. Since $A^2=A$, $Ax=x$. So $x-Bx=Cx$. By (2.2), $x-Bx=Cx=0$. Thus $F_B=\operatorname{Im} B$. By Proposition 2.7, $B^2=B$. Similarly $C^2=C$. Let $x \in R^n$. Since $A^2=A$, $BCx=-CBx$. By (2.2), $BCx=-CBx=0$. Hence $BC=CB=0$. ■

With Propositions 2.4 and 2.11, we can give shorter proofs for the following two propositions.

PROPOSITION 2.12. *Let $C \in M_{n\times n}$ such that $C^2=C^3$. Then C is an idempotent if and only if $r(C)=\operatorname{tr} C$ or $r(C)=r(C^2)$.*

PROPOSITION 2.13. *Let $A_1, A_2, \ldots, A_k$ be $n\times n$ matrices and let $A=\sum_{i=1}^k A_i$. Consider the following conditions:*

(a) *All A_i are idempotent.*
(b) *$A_iA_j=0$ for all $i\neq j$ and $r(A_i^2)=r(A_i)$, $i=1,2,\ldots,k$.*
(c) *A is an idempotent.*
(d) *$r(A)=\sum_{i=1}^k r(A_i)$.*

Then

(i) *any two of* (a), (b), *and* (c) *imply all* (a)–(d);
(ii) (c) *and* (d) *imply* (a) *and* (b).

The author is partially supported by NRC Grant A8518 and wishes to thank the referees for their helpful suggestions. Prop. 2.1, 2.4 were obtained at the Institute of Mathematics, Academia Sinica, R.O.C.

REFERENCES

1 T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, 1958.

2 K. S. Banerjee and G. Nagase, A note on the generalization of Cochran's theorem, *Comm. Statist. A—Theory Methods* 5 (9):837–842 (1976).

3 W. G. Cochran, The distribution of quadratic forms in a normal system with applications to the analysis of variance, *Proc. Cambridge Soc.* 30:178–191 (1934).

4 I. J. Good, Conditions for a quadratic form to have a chi square distribution, *Biometrika* 56:215–216 (1969).

5 K. Hoffman and R. Kunze, *Linear Algebra*, Prentice-Hall, 1961.

6 C. G. Khatri, Some results for the singular normal multivariate regression models, *Sankyā Ser. A* 30:267–280 (1968).

7 R. M. Loynes, On idempotent matrices, *Ann. Math. Stat.* 37:292–296 (1969).

8 G. Nagase and K. S. Banerjee, On the distribution of a second degree polynomial statistics in singular normal variates, *Comm. Statist. A—Theory Methods*, 5 (9):407–412 (1976).

9 S. R. Searle, *Linear Models*, Wiley, New York, 1971.

10 G. P. H. Styan, Notes on the distribution of quadratic forms in singular normal variables, *Biometrika* 57:567–572 (1970).

11 W. Y. Tan, Some matrix results and extensions of Cochran's theorem, *SIAM J. Appl. Math.* 28:547–554 (1975).

Received 28 January 1980; revised 13 May 1981