Canad. Math. Bull. Vol. **15** (1), 1972

ON THE SOLUTIONS OF THE MATRIX EQUATION $f(X, X^*) = g(X, X^*)$

BY
P. BASAVAPPA

It is well known that the matrix identities $XX^*=I$, $X=X^*$ and $XX^*=X^*X$, where X is a square matrix with complex elements, X^* is the conjugate transpose of X and I is the identity matrix, characterize unitary, hermitian and normal matrices respectively. These identities are special cases of more general equations of the form (a) $f(X, X^*)=A$ and (b) $f(X, X^*)=g(X, X^*)$, where $f(x, y)$ and $g(x, y)$ are monomials of one of the following four forms: $xyxy\ldots xyxy$, $xyxy\ldots xyx$, $yxyx\ldots yxyx$, and $yxyx\ldots yxy$. In this paper all equations of the form (a) and (b) are solved completely. It may be noted a particular case of $f(X, X^*)=A$, viz. $XX'=A$, where X is a real square matrix and X' is the transpose of X was solved by Weitzenböck [3]. The distinct equations given by (a) and (b) are enumerated and solved.

Most of the terminology is standard. All the matrices are matrices of complex numbers. By a projection is meant a matrix E such that $E=E^*=E^2$.

The main tools used in the solutions of the equations are: (1) the principal axis theorem for a nonhermitian matrix [1] and (2) the polar decomposition of a matrix [2]. These are stated as lemmas for later use.

LEMMA 1. *Let X be any rectangular matrix. Then there exist unitary matrices U and V such that*

$$UXV = \operatorname{diag}(x_1, \ldots, x_r, 0, \ldots, 0),$$

where $x_1, \ldots, x_r, 0, \ldots, 0$ are singular values of X.

LEMMA 2. *Let X be any square matrix. Then X can be written as*

$$X = HU(VK)$$

where $H(K)$ is p.s.d. and is unique and $U(V)$ is a unitary matrix. Moreover $H(K)$ and $U(V)$ commute if and only if X is normal.

THEOREM 1. *A matrix X is a solution of the matrix equation*

$$(XX^*)^p = A, \quad p \geq 1,$$

iff

$$X = U^* \operatorname{diag}(\alpha_1^{1/2p}, \ldots, \alpha_r^{1/2p}, 0, \ldots, 0)V^*$$

Received by the editors September 24, 1970 and, in revised form, October 21, 1970.

https://doi.org/10.4153/CMB-1972-010-9 Published online by Cambridge University Press

where U is any unitary matrix such that

$$UAU^* = \text{diag}(\alpha_1, \ldots, \alpha_r, 0, \ldots, 0)$$

and V is any unitary matrix.

Proof. Assume X is a solution of the equation. By Lemma 1 we have

$$UXV = \text{diag}(x_1, \ldots, x_r, 0, \ldots, 0). \tag{1}$$

Since X is a solution of the equation, we get

$$((UXV)(UXV)^*)^p = UAU^*. \tag{2}$$

Therefore if $\alpha_1, \ldots, \alpha_r, 0, \ldots, 0$ are the characteristic roots of A, by making use of (1) in (2) we get

$$\text{diag}(x_1^{2p}, \ldots, x_r^{2p}, 0, \ldots, 0) = \text{diag}(\alpha_1, \ldots, \alpha_r, 0, \ldots, 0)$$

which implies that

$$UXV = \text{diag}(\alpha_1^{1/2p}, \ldots, \alpha_r^{1/2p}, 0, \ldots, 0)$$

or

$$X = U^* \text{diag}(\alpha_1^{1/2p}, \ldots, \alpha_r^{1/2p}, 0, \ldots, 0)V^*.$$

It is easily verified X in the above form satisfies the equation. Note that if $p=1$ and X is a real square matrix, we get Weitzenböck's result.

THEOREM 2. *A matrix X is a solution of the matrix equation*

$$(XX^*)^p X = A$$

iff

$$X = U^* \text{diag}(\alpha_1^{1/(2p+1)}, \ldots, \alpha_r^{1/(2p+1)}, 0, \ldots, 0)V^*$$

where $\alpha_1, \ldots, \alpha_r, 0, \ldots, 0$ are the singular values of A and U and V are the unitary matrices such that

$$UAV = \text{diag}(\alpha_1, \ldots, \alpha_r, 0, \ldots, 0).$$

Proof. Let X be a solution of the equation. Let U and V be matrices as in Theorem 1. From the given equation, we get

$$((UXV)(UXV)^*)^p UXV = UAV.$$

or

$$\text{diag}(x_1^{2p+1}, \ldots, x_r^{2p+1}, 0, \ldots, 0) = UAV.$$

Thus if the singular values of A are $\alpha_1, \ldots, \alpha_r, 0, \ldots, 0$ we get

$$\text{diag}(x_1, \ldots, x_r, 0, \ldots, 0) = \text{diag}(\alpha_1^{1/(2p+1)}, \ldots, \alpha_r^{1/(2p+1)}, 0, \ldots, 0).$$

Thus by (1) we get

$$X = U^* \operatorname{diag}(\alpha_1^{1/(2p+1)}, \ldots, \alpha_r^{1/(2p+1)}, 0, \ldots, 0)V^*.$$

If X is in the above form, it follows easily that X satisfies the equation.

THEOREM 3. *A matrix X is a solution of the matrix equation*

$$(XX^*)^p = (XX^*)^q, \quad p > q \geq 1$$

iff each nonzero singular value of X is 1.

Proof. Assume X is a solution of the equation.

By making use of Lemma 1 and using the same method as in the proof of Theorem 1, we get

$$\operatorname{diag}(x_1^{2p}, \ldots, x_r^{2p}, 0, \ldots, 0) = \operatorname{diag}(x_1^{2q}, \ldots, x_r^{2q}, 0, \ldots, 0)$$

which implies $x_1 = \cdots = x_r = 1$.

Now suppose X is a matrix with each of its nonzero singular value 1. Then by Lemma 1, we have

$$UXV = \operatorname{diag}(1, \ldots, 1, 0, \ldots, 0).$$

or

$$X = U^* \operatorname{diag}(1, \ldots, 1, 0, \ldots, 0)V^*.$$

It is easily checked X in the above form satisfies the equation.

THEOREM 4. *A matrix X is a solution of the equation*

$$(XX^*)^p X = (XX^*)^q X, \quad p > q \geq 0$$

iff each nonzero singular value of X is 1.

Proof. The proof as in Theorem 3 works.

THEOREM 5. *A matrix X is a solution of the matrix equation*

$$(XX^*)^p = (X^*X)^q, \quad p, q \geq 1, \quad p \neq q$$

iff X is normal with each of its nonzero singular value 1.

Proof. Note here X must be square. Assume X is a solution of the equation. By Lemma 2, we can write

$$X = UH.$$

Thus from the given equation, we get

(3) $$UH^{2p}U^* = H^{2q}.$$

It follows that H^{2p} and H^{2q} have the same eigenvalues. If the eigenvalues of H are $\lambda_1 \geq \cdots \geq \lambda_r > \lambda_{r+1} = \cdots = \lambda_n = 0$, we get

$$\lambda_i^{2p} = \lambda_i^{2q}, \quad i = 1, \ldots, r.$$

Thus $\lambda_1 = \cdots = \lambda_r = 1$. Hence H is a projection and from (3) we see that H and U commute. Therefore by Lemma 2, H is normal. If X is normal with each of its nonzero singular value one, as in the proof of Theorem 3, it follows that X satisfies the equation.

THEOREM 6. *A matrix X is a solution of the matrix equation*

$$(XX^*)^p X = (XX^*)^q$$

iff X is a projection.

Proof. Assume X is a solution of the equation. Then by making use of Lemma 2, from the given equation, we get

$$H^{2p+1}U = H^{2q} = H^{2q}I \tag{4}$$

By the uniqueness of H, we get

$$H^{2p+1} = H^{2q}.$$

Thus H is a projection. From (4), we have

$$HU = H$$
$$X = HU = H$$

is a projection.

The converse is obvious.

THEOREM 7. *A matrix X is a solution of the matrix equation*

$$(XX^*)^p X = (X^*X)^p X^* (p \geq 0)$$

iff X is hermitian.

Proof. Assume X is a solution. Then

$$\begin{aligned}(XX^*)^{2p+1} &= ((XX^*)^p X)(X^*(XX^*)^p) \\ &= ((X^*X)^p X^*)(X(X^*X)^p) \\ &= (X^*X)^{2p+1}.\end{aligned}$$

Therefore by the uniqueness of root extraction we get

$$XX^* = X^*X.$$

Hence X is normal. Therefore each eigenvalue of X satisfies

$$|\lambda|^p\lambda = |\lambda|^p\bar{\lambda}.$$

It follows that nonzero eigenvalues of X are real. Hence X is hermitian. The converse is obvious.

THEOREM 8. *A matrix X is a solution of the matrix equation*

$$(XX^*)^p X = (X^*X)^q X^*, \quad p \neq q,$$

iff X is hermitian with 0, 1, -1 *as the only eigenvalues.*

Proof. Assume X is a solution of the equation. As in the proof of Theorem 7, we get

$$(XX^*)^{2p+1} = (X^*X)^{2q+1}. \tag{5}$$

Since the eigenvalues of XX^* and X^*X coincide, it follows that any eigenvalue of XX^* is 1 or 0.

Therefore XX^* and X^*X are projections. From (5) we see that

$$XX^* = X^*X. \tag{6}$$

From (6) X is normal, therefore the eigenvalues λ of X satisfy the same equation as X:

$$|\lambda|^p \lambda = |\lambda|^q \bar{\lambda}.$$

Since XX^* is a projection, $|\lambda| = 1$ or 0. These facts immediately combine to give $\lambda = 0$, 1, or -1. The converse is easily checked.

ACKNOWLEDGEMENT. The author is grateful to the referee for suggesting many alternative proofs. The author is also grateful to Professor Naoki Kimura for suggesting the problem and to Professor N. Shklov for his encouragement.

REFERENCES

1. C. Eckart, and G. Young, *A principal axis transformation for non-hermitian matrices*, Bull. Amer. Math. Soc. **45** (1939), 118–121.

2. F. R. Gantmacher, *The theory of matrices*, Vol. I (translated from Russian). Chelsea, New York, 1959.

3. R. Weitzenböck, *Über die Matrixgleichung $XX' = A$.*, Proc. Akad. Wet. Amsterdam, **35** (1932), 328–330.

UNIVERSITY OF NORTH CAROLINA AT CHARLOTTE,
CHARLOTTE, NORTH CAROLINA