THE LEAST SQUARE SOLUTION TO QUATERNION MATRIX EQUATIONS

Xue Youcai

Received March 11, 1999; revised February 13, 2001

Abstract. The paper gives the definition and the expressions of a norm of a least square solutions to the quaternion matrix equations $AXB = C$ and the equation with the constraint condition $DX = E$, either.

1 Introduction The real quaternion matrix equations

$$AXB = C \tag{1}$$

are very important and investigated deeply in reference [1-2], in [3], given expressions of the least square solutions of the quaternion matrix equation

$$AX = B \tag{2}$$

and the equation with the constraint condition

$$DX = E. \tag{3}$$

In the paper, we define a norm of a real quaternion matrix, give expressions of the least square solution of the quaternion matrix equation (1) and the equation with the constraint condition (3).

Throughout the paper, we denote the real quaternion field by Q, the set of all $m \times n$ matrices over Q by $Q^{m\times n}$, the real part of a real quaternion b by $Re(b)$, the conjugate transpose of a matrix A by A^*, the trace of A by trA, the Moore-Penrose inverse of a real quaternion matrix A by A^+ (See[4]).

Definition 1 Let V be a generalized unitary space (See[3]), $\alpha \in V$, then $\sqrt{(\alpha, \alpha)}$ is called the norm of α and denoted by $\|\alpha\|$.

It is easy to prove the following:

Lemma 1 Let V be a generalized unitary space, $\alpha, \beta \in V$, then
$\|\alpha + \beta\|^2 = \|\alpha\|^2 + \|\beta\|^2 + 2Re(\alpha, \beta)$.

Definition 2 Let $A \in Q^{m\times n}$, then $\sqrt{trA^*A}$ is called the norm of the matrix A and denoted by $\|A\|$.

2000 *Mathematics Subject Classification.* 15A24, 15A33.
Key words and phrases. the quaternion field; matrix equation; the least square solution; the least norm solution.

2 Main Results Consider matrix equation (1) with $A\in Q^{m\times n}, B\in Q^{r\times t}, C\in Q^{m\times t}$.

Definition 3 For (1), if there exists $X_0\in Q^{n\times r}$ such that

$$\|AX_0B-C\|^2=\min_{X\in Q^{n\times r}}\|AXB-C\|^2 \tag{4}$$

then X_0 is called the least square solution of (1).

Definition 4 The matrix equation

$$A^*AXBB^*=A^*CB^* \tag{5}$$

is called the normal equation of (1).

Lemma 2 The matrix equation (1) is consistent if and only if

$$AA^+CB^+B=C. \tag{6}$$

In which case the general solution of (1) is

$$X=A^+CB^++Y-A^+AYBB^+, \tag{7}$$

where $Y\in Q^{n\times r}$ is arbitrary.

Lemma 3 (see[5]) Let $A\in Q^{m\times n}$, then
(*i*) $A^+=(A^*A)^+A^*=A^*(AA^*)^+$;
(*ii*) $A^*=A^*A(A^*A)^+A^*=A^*(AA^*)^+(AA^*)$

Lemma 4 (see[6]) Let $A\in Q^{m\times n}, B\in Q^{n\times m}$, then $Re(trAB)=Re(trBA)$.

Lemma 5 Let $L\in Q^{m\times n}, Q_L=I-L^+L, G\in Q^{m\times n}, N\in Q^{n\times n}$, then
(i) $Q_L^*=Q_L=Q_L^2$;
(ii) $Q_L(L^+G)=0$;
(iii) $Q_L(NQ_L)^+=(NQ_L)^+$.
It is easy to prove the follwing by lemma 2 and lemma 3.

Lemma 6 The normal equation (5) of the matrix equation (1) is consistent.

Lemma 7 (see[3]) The matrix equation (2) is consistent if and only if $AA^+B=B$. In which case the general solution of (2) is $X=A^+B+(I-A^+A)Y$, where $Y\in Q^{n\times r}$ is arbitrary.

Theorem 1 X_0 is the least square solution of (1) if and only if X_0 is a solution of (5), whence (1) has the least square solution.

Proof. By lemma 1 and lemma 4, for arbitrary $X, X_1\in Q^{n\times r}$,

$$\|AXB-C\|^2=\|A(X_1+X-X_1)B-C\|^2=\|AX_1B-C\|^2+\|A(X-X_1)B\|^2$$
$$+2Re[tr(X-X_1)^*A^*(AX_1B-C)B^*]. \tag{8}$$

Suppose X_0 is a solution of (5) i.e

$$A^*AX_0BB^*=A^*CB^*. \tag{9}$$

Let $X_1=X_0$ in (8), then by (9)
$\|AXB-C\|^2=\|AX_0B-C\|^2+\|A(X-X_0)B\|^2\geq\|AX_0B-C\|^2$.

Hence it follows from $X \in Q^{n\times r}$ is arbitrary that $\min_X \|AXB-C\|^2 \geq \|AX_0B-C\|^2$.
Consequently

$$\|AX_0B-C\|^2 = \min_X \|AXB-C\|^2, \quad (10)$$

i.e., X_0 is one of the least square solutions of (1).
Conversely, suppose X_0 is one of the least square solutions of (1) , then (10) holds. We may assume Y_0 is a solution of (5), i.e. $A^*AY_0BB^* = A^*CB^*$, by imitating the proof of (10),

$$\|AY_0B-C\|^2 = \min_X \|AXB-C\|^2. \quad (11)$$

By (10) and (11)

$$\|AX_0B-C\|^2 = \|AY_0B-C\|^2. \quad (12)$$

In (8), let $X = X_0, X_1 = Y_0$, then by (9)

$$\|AX_0B-C\|^2 = \|AY_0B-C\|^2 + \|A(X_0-Y_0)B\|^2.$$

Hence by (12), $\|A(X_0-Y_0)B\|^2 = 0$. Accordingly, $AX_0B = AY_0B$, whence $A^*AX_0BB^* = A^*AY_0BB^* = A^*CB^*$. This implies that X_0 is a solution of (5). By lemma 6, (1) has the least square solution.

Theorem 2 The set of the least square solutions of the matrix equation (1) is

$$M = \{A^+CB^+ + Y - A^+AYBB^+ | Y \in Q^{n\times r}\}. \quad (13)$$

Proof. By theorem 1, we only need to prove that the set of solutions of (5) can be expressed as (13). Since (5) is solvable, by lemma 2, the set of solutions of (5) is

$$\{(A^*A)^+A^*CB^*(BB^*)^+ + Y - (A^*A)^+(A^*A)Y(BB^*)(BB^*)^+ | Y \in Q^{n\times r}\}. \quad (14)$$

By lemma 3, (14) is (13).

Definiton 5 Let $X_0 \in M$, if $\|X_0\| = \min_{X\in M} \|X\|$, then X_0 is called the least norm solution of (1) .

Theorem 3 A^+CB^+ is the least norm solution of (1).

Proof. By theorem 2, $A^+CB^+ \in M$. Suppose X_1 is a least square solution of (1), i.e. $X_1 = A^+CB^+ + Y_1 - A^+AY_1BB^+, Y_1 \in Q^{n\times r}$, then

$$\|X_1\|^2 = \|A^+CB^+\|^2 + \|Y_1 - A^+AY_1BB^+\|^2 + 2Re[tr(Y_1 - A^+AY_1BB^+)^*A^+CB^+]. \quad (15)$$

By lemma 4,

$$\begin{aligned}
& Re[tr(Y_1 - A^+AY_1BB^+)^*A^+CB^+] \\
= \;& Retr[(Y_1{}^* - BB^+Y_1{}^*A^+A)A^+CB^+] \\
= \;& Re[tr(Y_1{}^*A^+CB^+ - BB^+Y_1{}^*A^+AA^+CB^+)] \\
= \;& Re[tr(Y_1{}^*A^+CB^+ - BB^+Y_1{}^*A^+CB^+)] \\
= \;& Re[tr(I - BB^+)Y_1{}^*A^+CB^+] \\
= \;& Re[tr(Y_1{}^*A^+CB^+(I - BB^+))] \\
= \;& Re[tr(Y_1{}^*A^+CB^+ - Y_1{}^*A^+CB^+BB^+] \\
= \;& Re[tr(Y_1{}^*A^+CB^+ - Y_1{}^*A^+CB^+)] = 0.
\end{aligned}$$

By (15), $\|X_1\|^2 = \|A^+CB^+\|^2 + \|Y_1 - A^+AY_1BB^+\|^2 \geq \|A^+CB^+\|^2$, i.e., A^+CB^+ is one of the least norm solution of (1).

Corollary 1 (see[3]) A^+B is the least norm solution of (2).

Definition 6 For (1) with $DX = E, X_0 \in Q^{n\times r}$, is called one of the least square solution of the matrix equation (1) with constraint condition (3) if X_0 satisfies

i) $DX_0 = E$;

ii) $\|AX_0B - C\|^2 = \min\limits_{DX=E} \|AXB - C\|^2$.

Theorem 4 Let $DD^+E = E$, then the set of the least square solution of the matrix equation (1) with the constraint condition (3) is

$$(16)\quad S = \{D^+E + Q_D(AQ_D)^+(C - AD^+EB)B^+ + Q_DV - Q_D(AQ_D)^+(AQ_D)VBB^+\},$$

where $Q_D = I - D^+D, V \in Q^{n\times r}$.

Proof. By lemma 7, the equation (3) is consistant and the general solution is

$$(17)\qquad X = D^+E + Q_DU,$$

where $U \in Q^{n\times r}$ is arbitrary. Hence

$$(18)\qquad \min_{DX=E} \|AXB - C\|^2 = \min_{U} \|AQ_DUB - (C - AD^+EB)\|^2.$$

By theorem 2 and lemma 3.

$$(19)\qquad U = (AQ_D)^+(C - AD^+EB)B^+ + V - (AQ_D)^+(AQ_D)VBB^+,$$

where, $V \in Q^{n\times r}$ is arbitrary. (16) follows from is inserting (19) into (17).

Corollary 2 (see[3]) Let $DD^+E = E$, then the set of the least square solutions of (2) with the constraint condition (3) is $T = \{D^+E + Q_D(AQ_D)^+(B - AD^+E) + Q_DZ - Q_D(AQ_D)^+(AQ_D)Z; Z \in Q^{n\times n}\}$.

Definition 7 Let $X_0 \in S$, if $\|X_0\|^2 = \min\limits_{X\in S} \|X\|^2$, then X_0 is called the least norm solution of (1) with the constraint condition (3).

Theorem 5 Suppose $DD^+E = E$, then

$$(20)\qquad X_0 = D^+E + Q_D(AQ_D)^+(C - AD^+EB)B^+$$

is the least norm solution of (1) with the constraint condition (3).

Proof. For $X_1 \in S$ is arbitrary, then

$$\|X_1\|^2 = \|Q_DV_1 - Q_D(AQ_D)^+(AQ_D)V_1BB^+\|^2 + \|D^+E + Q_D(AQ_D) + (C - AD^+EB)B^+\|^2$$

$$(21)\quad + 2Re[tr(Q_DV_1 - Q_D(AQ_D)^+(AQ_D)V_1BB^+)^*(D^+E + Q_D(AQ_D)^+(C - AD^+EB)B^+)]$$

where, $V_1 \in Q^{n\times r}$. By lemma 4 and lemma 5, $Re[tr(Q_DV_1 - Q_D(AQ_D)^+(AQ_D)V_1BB^+)^*(D^+E + Q_D(AQ_D)^+(C - AD^+EB)B^+)] = 0$, then $\|X_1\|^2 \geq \|D^+E + Q_D(AQ_D)^+(C - AD^+EB)B^+\|^2 = \|X_0\|^2$, i.e., X_0 is the least norm solution of (1) with the constraint condition (3).

Similarly we can consider the quaternion matrix equation

$$(22)\qquad AXA^* = B.$$

Likewise we can define the least square solution and the least norm solution of (22) with the constraint condition (3) and we have the similar results like theorem 1-5.

References

[1] Wang Qingwen, *Metapositive semidefinite solutions to the matrix equation $AXB = C$ over a strong p–division ring,* Northeast Math.J. (PRC), 12:2(1996).

[2] Xue Youcai and Wang Qingwen, *Minor self–conjugate solutions of Linear Matrix Equations over a skew field,* Banyan Math.J. 4 (1997), 21–30.

[3] Xue Youcai, *The least square solutions to the Quaternion Matrix Equation $AX = B$,* Journal of Math. 17:1(1997), 87–90.

[4] Zhuang Wajin, *Generalized inverses of matries over a skew field,* J. of Maths. (PRC), 6:1(1986), 105–112.

[5] Zhang Shuqing, *Some Properties of generalized over the p–division ring.* J. of Qufu Teachers University (ShanDong), 18:3(1992), 33–36.

[6] Zhang Shuqing, *Some theorems of traces of quaterninon matrices,* J. of Math. Res and Exp., 13:(1993), 567–572.

DEPARTMENT OF MATHEMATICS,
YUNCHENG ADVANCED TRAINING COLLEGE,
YUNCHENG SHANXI, 044000 P.R.CHINA