International Journal of Pure and Applied Mathematics
Volume 24 No. 1 2005, 131-137

ON THE SOLUTION OF SYLVESTER, LYAPUNOV AND STEIN EQUATIONS OVER ARBITRARY RINGS

M.C. Gouveia
Department of Mathematics
University of Coimbra
Largo D. Dinis, Coimbra, 3000, PORTUGAL
e-mail: mcag@mat.uc.pt

Abstract: The theory of generalized inverses is used to discuss the consistency of each of the matrix equations $AX-YB=C$, $AX-XB=C$ and $AXB-X=C$, over a noncommutative ring. Some explicit general solutions are also presented.

AMS Subject Classification: 15A09
Key Words: Sylvester, Lyapunov, Stein equations, generalized inverses

1. Introduction

In 1952 Roth stated in his fundamental paper [9] that, over a field, the Sylvester equation $AX-YB=C$ has a solution if and only if the matrices

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix}, \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$$

are equivalent.

Also the Lyapunov equation $AX-XB=C$ can be solved if and only if the foregoing matrices are similar.

Roth's proofs were based on canonical forms and other proofs have been given using linear transformations and dimension arguments, Flanders et al [3]. Analogous results for other linear matrix equations such as Stein equation $X-AXB=C$ were established at Wimmer [10]. Roth's Theorems have also been proven for matrices over commutative, Gustafson [4], and unit regular rings,

Received: September 18, 2005

Hartwig [5], but most proofs are existence proofs not furnishing any solution. J. Baksalary and R. Kala [1] established the consistency and gave explicit solutions to the Sylvester equation taking into account that every matrix over a field has a generalized inverse. Those results were discussed in Huylebrouck [7] for regular matrices over noncommutative rings.

In this work we want to deal with consistency and explicit solutions for those matrix equations over an arbitrary ring without multiplicative identity. We will call it a *rng*, Jacobson [8], and we will denote it by R. If R has a multiplicative identity we call it a *ring with unity*.

Suppose $\phi : X \to Y$ is a morphism of a category $\mathcal{C}$. Then ϕ is said to be *von Neumann regular* if there is a morphism $\chi : Y \to X$ of $\mathcal{C}$ such that $\phi\chi\phi = \phi$. χ is called a (1)-*inverse* of ϕ and we will denote it by $\phi^{(1)}$. In this case $\phi\ \phi^{(1)}$and $\phi^{(1)}\phi$ are idempotent endomorphisms of Y and X, respectively. If χ also satisfies the equation $\chi\phi\chi = \chi$ then it is called a (1,2)-*inverse* of ϕ . We note that if $\phi^{(1)}$exists, then we always may consider $\phi^{(1)}\phi\phi^{(1)}$as a (1,2)-inverse of ϕ. Moreover, if χ is a (1,2)-inverse which satisfies the equation $\phi\chi = \chi\phi$ then $\chi = \phi^{\#}$ is a *group inverse* of ϕ . If it exists, the group inverse is unique, Ben-Israel et al [2].

If we let $\mathrm{Mat}(R)$ be the additive category of finite matrices with elements in R and if for any positive integers m, n we identify $\mathrm{Hom}_R(nR, mR)$ with the set $M_{m,n}(R)$ of all $m \times n$-matrices over R (in the standard manner) then the following results also apply to morphisms in additive categories.

2. Main Results

A relation which is weaker than rank equivalence and similarity is the so-called $g - relation$, Hartwig et al [6] which is defined by

$$A - g - B \text{ if and only if } PAQ = B,\ RBS = A \text{ for some } P, Q, R, S. \tag{1}$$

Stronger than g-relation but still weaker than similarity is *pseudosimilarity*, Hartwig et al[6]. For elements A, B in $\mathrm{Mat}(R)$, we say that *A is pseudosimilar to B if and only if*

$$P^{(1)}AP = B, PBP^{(1)} = A \text{ for some von Neumann regular } P. \tag{2}$$

Now, if in (1)

$$R = P^{(1)} \text{ and } S = Q^{(1)} \text{ we say that } A \text{ and } B \text{ are pseudoequivalent.} \tag{3}$$

We note that, as it is known, the g-relation coincides with rank equivalence over a division ring and, over a field, pseudosimilarity and similarity also coincide.

Our aim is to investigate if Roth's equivalence and similarity relations can be somehow extended to general rngs while the concept of generalized inverses is required.

Theorem 1. *Let R be a noncommutative rng. Let $A \in M_m(R), B \in M_n(R)$ be group invertible matrices and $C \in M_{m,n}(R)$ arbitrary. Let*

$$M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}, \quad D = \mathrm{diag}(A, B).$$

Then the following are equivalent.

(i) There is a matrix $X \in M_{m,n}(R)$ such that

$$AX - XB = C.$$

(ii) There is a von Neumann regular matrix P such that M and D are pseudosimilar and such that

$$D = DP^{(1)}P = P^{(1)}PD.$$

Proof. Since A, B are group invertible, if $AX - XB = C$ is consistent then it is clear that there are matrices

$$P = \begin{pmatrix} AA^{\#} & -XBB^{\#} \\ 0 & BB^{\#} \end{pmatrix}, \quad P^{(1)} = \begin{pmatrix} A^{\#}A & A^{\#}AX \\ 0 & B^{\#}B \end{pmatrix}$$

such that $M = PDP^{(1)}$ and $P^{(1)}MP = D$.

Moreover, $D = DP^{(1)}P = P^{(1)}PD$. □

Now, if $P, P^{(1)}$ exist such that M and D are pseudosimilar and $D = DP^{(1)}P = P^{(1)}PD$, then M is von Neumann regular with a (1,2)-inverse given by

$$M^{(1)} = P \begin{pmatrix} A^{\#} & 0 \\ 0 & B^{\#} \end{pmatrix} P^{(1)},$$

and

$$MP - PD = 0,\ M^{(1)}P - PD^{(1)} = 0 \tag{4}$$

are consistent equations.

Moreover, if R is a ring with unity, then M and $M^{(1)}$ belong to the ring with unity

$$\Re = \left\{ \begin{pmatrix} U & T \\ 0 & V \end{pmatrix} : U \in M_m(R), V \in M_n(R), T \in M_{m,n}(R) \right\},$$

and $M = D + J$, where D is von Neumann regular and $J = \begin{pmatrix} 0 & C \\ 0 & 0 \end{pmatrix}$ belongs to the Jacobson radical of $\Re R$, $\mathrm{rad}(\Re R)$. It follows that, Huylebrouck [7],

$$AA^{\#}C - AA^{\#}CB^{\#}B + CB^{\#}B = C. \tag{5}$$

Otherwise, if R is a ring without unity, let be the natural imbedding of R into a ring R^* with unity. Then $f(M)$ belongs to the ring with unity

$$\Re^* = \left\{ \begin{pmatrix} f(U) & f(T) \\ 0 & f(V) \end{pmatrix} : f(U) \in M_m(R^*), f(V) \in M_n(R^*), f(T) \in M_{m,n}(R^*) \right\}.$$

Since f is a rng homomorphism, $f(M)$ is von Neumann regular with $f(M^{(1)}) = (f(M))^{(1)}$, and $f(M) = f(D) + f(J)$, where $f(D)^{(1)}$ =diag $(f(A^{(1)}), f(B)^{(1)})$ and $f(J) \in \mathrm{rad}(\Re^*)$.

Hence, equality (5) holds in $\mathrm{Mat}(R^*)$ for the images of the blocks of M and $M^{(1)}$. Taking into account that f is a monic morphism from R onto $f(R)$, then (5) holds in $\mathrm{Mat}(R)$.

Finally, from (4) and (5) we have $P \in \mathrm{Ker}\phi$, with $\phi : K \to MK - KD$, is of the form

$$P = \begin{pmatrix} A^{\#}A & S \\ 0 & B^{\#}B \end{pmatrix},$$

where S is a solution of $AX - XB = C$.

Remark 2. We notice that if in Theorem 1, A and B are just von Neumann regular then (i) implies that M *and* D are pseudoequivalent. In fact the hypotheses in (3) are satisfied for von Neumann regular

$$P = \begin{pmatrix} AA^{(1)} & -XBB^{(1)} \\ 0 & BB^{(1)} \end{pmatrix}, \qquad Q = \begin{pmatrix} A^{(1)}A & A^{(1)}AX \\ 0 & B^{(1)}B \end{pmatrix}.$$

Theorem 3. *Let R be a noncommutative rng. Let $A \in M_{m,r}(R), B \in M_{s,n}(R)$ be group invertible matrices and $C \in M_{m,n}(R)$ arbitrary. Let*

$$M = \begin{pmatrix} A & C \\ 0 & B \end{pmatrix}, \quad D = \mathrm{diag}(A, B).$$

Then the following are equivalent.

(i) There are matrices $X \in M_{r,n}(R)$ *and* $Y \in M_{m,s}(R)$ *such that*

$$AX - YB = C.$$

(ii) The matrices M *and* D *are pseudoequivalent .*

(iii)

$$AA^{(1)}C - AA^{(1)}CB^{(1)}B + CB^{(1)}B = C.$$

In that case, the general solution of $AX - YB = C$ *has the form*

$$\begin{aligned} X &= A^{(1)}C + A^{(1)}ZB + W - A^{(1)}AW, \\ Y &= -CB^{(1)} + AA^{(1)}CB^{(1)} + Z + AA^{(1)}ZBB^{(1)} - ZBB^{(1)}, \end{aligned} \tag{6}$$

with $W \in M_{r,n}(R)$ *and* $Z \in M_{m,s}(R)$ *being arbitrary.*

Proof. Well known methods yield $(i) \Longleftrightarrow (iii)$, (see Huylebrouck [7]).

(ii)$\Longleftrightarrow$(iii) If P, Q, are von Neumann regular matrices such that $M = PDQ$ and $D = P^{(1)}MQ^{(1)}$ then M is von Neumann regular. Indeed, $M^{(1)} = Q^{(1)}D^{(1)}P^{(1)}$. Hence the rest of the proof follows using a similar reasoning to that of the proof of Corollary 5 in Huylebrouck [7].

Now, if $AA^{(1)}C - AA^{(1)}CB^{(1)}B + CB^{(1)}B = C$, then there are

$$P = \begin{pmatrix} AA^{(1)} & -AA^{(1)}CB^{(1)} + CB^{(1)} \\ 0 & BB^{(1)} \end{pmatrix}, \quad Q = \begin{pmatrix} A^{(1)}A & A^{(1)}C \\ 0 & B^{(1)}B \end{pmatrix}$$

and

$$P^{(1)} = \begin{pmatrix} AA^{(1)} & 0 \\ 0 & BB^{(1)} \end{pmatrix}, \quad Q^{(1)} = \begin{pmatrix} A^{(1)}A & -A^{(1)}CB^{(1)}B \\ 0 & B^{(1)}B \end{pmatrix}$$

such that M and D are pseudoequivalent.

Finally, it is clear that $X = A^{(1)}C$ and $Y = -CB^{(1)} + AA^{(1)}CB^{(1)}$ is a particular solution of the Sylvester equation. An easy computation leads to (6) as a general solution in $\mathrm{Mat}(R)$.

Corollary 4. *Let* $A \in M_m(R), B \in M_n(R)$ *be von Neumann regular matrices. Let* $C \in M_{m,n}(R)$ *be arbitrary. Then there exists a solution* $Y \in M_{m,n}(R)$ *for the Stein equation*

$$AYB - Y = C \tag{7}$$

if and only if $A^{(1)}C - AA^{(1)}CB + CB = 0$. *In that case, the general solution is given by*

$$Y = -C + AA^{(1)}C + AA^{(1)}Z$$

with $Z \in M_{m,n}(R)$ *arbitrary.*

Proof. We write (7) as a set of two equations

$$AX - Y = C\,, \quad X - YB = 0\,. \tag{8}$$

It is clear that the second equation is consistent with solution $Y = Z$, $X = ZB$, Z arbitrary. The first equation is of Sylvester's type and, from Theorem 2, we can check that it is consistent with general solution

$$\begin{aligned} X &= A^{(1)}C + A^{(1)}Z + W - A^{(1)}AW\,, \\ Y &= -C + AA^{(1)}C + AA^{(1)}Z\,, \end{aligned} \tag{9}$$

where W and Z are arbitrary matrices.

Hence (9) is a solution of (8) if and only if $A^{(1)}C - AA^{(1)}CB + CB = 0$.

Remark 5. We notice that the explicit expressions for the solutions (6) in Baksalary et al [1] require a concept of (1)-inverse. Thus, for a generalization to noncommutative rings , regular rings or semi-simple Artinian rings are examples of natural candidates.

Remark 6. It is well known from the theory of generalized inverses that over a field the equation $AX = C$ is consistent if and only if $AA^{(1)}C = C$ (see Ben-Israel et al [2]). In that case a general solution is given by $X = A^{(1)}C + (I - A^{(1)}A)W$, W arbitrary. Hence, for $B = O$ and A von Neumann regular, Theorem 2 extends this result to noncommutative rngs.

Acknowledgements

This research was supported by IT – Institute of Telecommunications, Portugal.

References

[1] J.K. Baksalary, R. Kala, The matrix equation $AX - YB = C$, *Linear Algebra and its Appl.*, **25** (1979), 41-43.

[2] A. Ben-Israel, T.N.E. Greville, *Generalized Inverses: Theory and Applications*, John Wiley and Sons, Inc. (1974).

[3] H. Flanders, H.K. Wimmer, On the matrix equations $AX - XB = C$ and $AX - YB = C$, *SIAM J. Appl. Math.*, **32**, No. 4 (1977), 707-710.

[4] W. Gustafson, Roth's Theorems over commutative rings, *Linear Algebra and its Appl.*, **23** (1979), 245-251.

[5] R.E. Hartwig, Roth's equivalence problem in unit regular rings, *Proc. Amer. Math. Soc.*, **59**, No. 1 (1976), 39-44.

[6] R.E. Hartwig, F. Hall, Pseudo-similarity for matrices over a field, *Proc. Amer. Math. Soc.*, **71** (1978), 6-10.

[7] D. Huylebrouck, The generalized inverse of a sum with radical element: applications, *Linear Algebra and its Appl.*, **246** (1996), 159-175.

[8] I. Jacobson, *Basic Algebra* , W.A. Freeman and Company (1985).

[9] W.E. Roth, The equations $AX - YB = C$ and $AX - XB = C$ in matrices, *Proc. Amer. Math. Soc.*, **3** (1952), 392-396.

[10] H.K. Wimmer, The matrix equation $X - AXB = C$ and an analogue of Roth's Theorem, *Linear Algebra and its Appl.*