BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **8**, No 2

Sofia • 2008

On the Sensitivity of the Matrix Equation $XA - AX = X^2$ *

Mihail Konstantinov[1], Petko Petkov[2], Ivan Popchev[3], Vera Angelova[3]

[1]University of Architecture, Civil Engineering and Geodesy, 1046 Sofia
E-mail: mmk_fte@uacg.bg
[2]Department of Automatics, Technical University of Sofia, 1756 Sofia
E-mail: php@tu-sofia.acad.bg
[3]Institute of Information Technologies, 1113 Sofia
E-mails: {ipopchev, vangelova}@iit.bas.bg

Abstract: *The paper deals with the matrix equation $XA - AX = X^2$ arising in the analysis of affine structures on solvable Lie algebras. The sensitivity of the equation relative to perturbations in the coefficient matrix A is studied. Both local and non-local perturbation bounds are obtained. Illustrative numerical examples demonstrate the effectiveness of the bounds proposed.*

Keywords: *perturbation analysis, matrix equations, Lie algebras, algebraic Riccati equations.*

Mathematics Subject Classification 2000: *15A24*

1. Introduction and notation

The matrix equation $XA - AX = X^p$ in X, where p is a positive integer and A, X are $n\times n$ matrices over an algebraically closed field $\mathbb{K}$ of characteristic zero, is connected with problems in Lie theory [1, 2]. The case when $p = 2$ arises in studying affine structures on solvable Lie algebras and is a special case of the algebraic Riccati equation. Further on we assume that $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

For any given matrix A the equation $XA - AX = X^2$ always has a solution, namely the trivial solution $X = 0$. If A has multiple eigenvalues then this equation

* This research work was supported by the Bulgarian Academy of Sciences under Grant No 010077/2007.

has non-trivial solutions. A special set of solutions is obtained for $XA - AX = 0,\ X^2 = 0$.

According to [1], for $p \geq 2$ every solution X of the equation $XA - AX = X^p$ is a nilpotent matrix and if A has no multiple eigenvalues then $X = 0$ is the only matrix solution to $XA - AX = X^p$. Conversely, if A has multiple eigenvalues then there exist nontrivial solutions. We also note that adding a scalar matrix to A does not change the form of the equation.

In this paper local and non-local perturbation bounds for the solution to the equation

(1) $$XA - AX = X^2,\ A, X \in \mathbb{K}^{n\times n},$$

are derived, where $\mathbb{K}^{n\times n}$ is the space of $n\times n$ matrices over $\mathbb{K}$.

Throughout the paper the following notations are used: $\mathbb{R}$ and $\mathbb{C}$ – the sets of real and complex numbers, respectively; I_n – the identity $n\times n$ matrix; vec(A) $\in \mathbb{K}^{n^2}$ – the column-wise vector representation of the matrix $A \in \mathbb{K}^{n\times n}$, where $\mathbb{K}^m = \mathbb{K}^{m\times 1}$; Mat($\mathcal{L}$) $\in \mathbb{K}^{n^2\times n^2}$ – the matrix representation of the linear matrix operator $\mathcal{L}$: $\mathbb{K}^{n\times n} \to \mathbb{K}^{n\times n}$; N_m – the $m\times m$ nilpotent matrix with elements $N_m(k, l) = 1$ for $l = k+1$, $k = 1, 2, \ldots, m-1$, and $N_m(k, l) = 0$ otherwise; $A \otimes B = [A(k, l)B]$ – the Kronecker product of the matrices $A = [A(k, l)]$ and B; $\|\cdot\|$ – a vector or a matrix norm; $\|\cdot\|_2$ – the Euclidean vector or the spectral matrix norm; $\|\cdot\|_F$ – the Frobenius norm.

The notation ':=' stands for 'equal by definition'.

2. Statement of the problem

Equation (1) may be written in the equivalent form

(2) $$F(X, A) := XA - AX - X^2 = 0,\ A, X \in \mathbb{K}^{n\times n}.$$

Denote by $S_A \subset \mathbb{K}^{n\times n}$ the set of all solutions to equation (2). As mentioned above, the set S_A is invariant relative to scalar shifts in A, i.e. $S_A = S_{A+\mu I_n}$ for all $\mu \in \mathbb{K}$.

We shall suppose that the following assumption holds true.

Assumption A1. The matrix A has multiple eigenvalues and equation (1) has non-trivial solutions, i.e. $S_A \neq \{0\}$.

Since every solution X is a nilpotent matrix [1] we have $X^p = 0$. So the interesting case is $p > 2$ since for $p = 2$ the equation reduces to the system $XA - AX = 0, X^2 = 0$, considered below.

Example 1. Let $n = 2$ and $\mathbb{K} = \mathbb{C}$. Since A has a double eigenvalue and the matrices A and $A + \mu I_2$, $\mu \in \mathbb{C}$, produce the same solution set S_A, we actually have the following two cases.

1. The first case is $A = 0$ and the system is reduced to equation $X^2 = 0$. Here the solution set S_0 is the union of an one-parametric variety $\{xN_2 : x \in \mathbb{C}\}$, and a

two-parametric family of solutions X with $X(1,1) = x \in \mathbb{C}$, $X(2, 1) = y \in \mathbb{C}$, $y \neq 0$, and $X(1, 2) = -x^2/y$, $X(2, 2) = -x$.

2. The second case is $A = N_2$. Here the solution set S_{N_2} is $\{xN_2 : x \in \mathbb{C}\}$. ■

Let the matrix A be subject to a perturbation E, so that the coefficient matrix becomes $A + E$. We shall consider only perturbations E from an admissible set $\mathcal{E} \subset \mathbb{K}^{n\times n}$ which satisfies the following assumptions.

Assumption A2. The matrix A is non-zero and the norm of the matrices from $\mathcal{E}$ is small compared to the norm of A.

Assumption A3. The perturbed equation

$$F(Y, A + E) = 0 \tag{3}$$

if Y has non-trivial solutions for all $E \in \mathcal{E}$, i.e. $S_{A+E} \neq \{0\}$ for $E \in \mathcal{E}$.

Denote any solution Y of (3) as $Y = X + Z$, where Z is a perturbation (not necessarily small) of a fixed solution $X_0 \in S_A$ of equation (1).

We recall that both equations (1) and (3) have multi-parametric families of solutions S_A and S_{A+E}, respectively. This means that for $X_0 \in S_A$ fixed we shall have a family

$$\mathcal{Z} = \mathcal{Z}(A, X_0) := \{Y - X_0 : Y \in S_{A+E}\} \subset \mathbb{K}^{n\times n}$$

of perturbations Z in X_0.

We stress that the sets S_A, S_{A+E} and $\mathcal{Z}$ may not be bounded. So we may not estimate the norm of any element of $\mathcal{Z}$ (by a function of the norm of E). Rather, we shall estimate the norm of certain elements $Z \in \mathcal{Z}$. In any case our bounds will estimate the quantity $\inf\{\|Z\|_F : Z \in \mathcal{Z}\}$ from above.

Our next assumption concerning the set $\mathcal{E}$ of admissible perturbations E in A is as follows.

Assumption A4. For any $\eta > 0$ there is $\delta = \delta(\eta) > 0$ such that there exists $Z \in \mathcal{Z}$ with $\|Z\| < \eta$ provided $E \in \mathcal{E}$ and $\|E\| < \delta$.

It is worth mentioning that Assumptions **A2**, **A3** and **A4** will be fulfilled for a set of small perturbations preserving the Jordan form of A.

Example 2. Let $n = 2$, $\mathbb{K} = \mathbb{C}$, $A = N_2$ and $\mathcal{E} \subset \mathbb{C}^{2\times 2}$ be the set of matrices $E = x N_2$ with $|x| < 1$. Then the perturbations $A \rightarrow A + E$ preserve the Jordan form of A and

$$S_A = S_{A+E} = \mathcal{Z} = \{xN_2 : :x \in \mathbb{C}\}.$$ ■

An important problem in studying perturbed equations of type (3) is to find local and non-local bounds for $\|Z\|_F$ as functions of the norm $\|E\|_F$ of the perturbation E in the data matrix A, where $Y = X_0 + Z$ and $X_0 \in S_A$. The local bound should be valid for $\|E\|_F$ asymptotically small, while the non-local bound will hold true for perturbations in the data belonging to a certain finite set containing the origin.

However, this general program may not be fulfilled completely since the standard technique of perturbation analysis [3] is not applicable to the problem considered. Rather, we shall obtain local bounds on the norm of certain projections of the perturbation Z on subspaces of $\mathbb{K}^{n\times n}$ of positive codimension.

3. Local perturbation analysis

Consider for simplicity the case $\mathbb{K} = \mathbb{R}$. The case $\mathbb{K} = \mathbb{C}$ is treated similarly.

Let $X_0 \in S_A$. Denote by $F_U(.) = F_U(X_0, A)(.)$ the partial Fréchet derivative of the function $F(.,.)$ in the argument $U \in \{X, A\}$ computed at the point (X_0, A) and define the operators

$$\mathcal{L}(.) := F_X(X_0, A)(.), \quad \mathcal{M}(.) := F_A(X_0, A)(.).$$

These are linear operators $\mathbb{K}^{n\times n} \to \mathbb{K}^{n\times n}$ such that

$$F(X_0 + Z, A + E) = F(X_0, A) + \mathcal{L}(Z) + \mathcal{M}(E) + G(Z, E), \tag{4}$$

and their action is given by

$$\mathcal{L}(H) = H(A - X_0) - (A + X_0)H, \tag{5}$$
$$\mathcal{M}(H) = X_0 H - H X_0, \quad H \in \mathbb{R}^{n\times n}.$$

The term $G(Z, E)$ contains the second order terms in Z, E,

$$G(Z, E) = Z E - E Z - Z^2 = O(u^2), \quad u \to 0,$$

where

$$u := \varepsilon + \|Z\|_F, \ \varepsilon := \|E\|_F.$$

In what follows it is supposed that the asymptotic estimates of the form $O(u^k)$ k=1, 2, are valid for $u \to 0$.

The matrix representations $L, M \in \mathbb{R}^{n^2\times n^2}$ of the operators $\mathcal{L}$, $\mathcal{M}$ are

$$\begin{aligned} L &:= (A - X_0)^{\mathrm{T}} \otimes I_n - I_n \otimes (A + X_0), \\ M &:= I_n \otimes X_0 - X_0^{\mathrm{T}} \otimes I_n. \end{aligned} \tag{6}$$

If the operator $\mathcal{L}$ is invertible, i.e. if its matrix representation L is non-singular, then the perturbed equation (3) may be rewritten as an equivalent matrix equation [2, 3, 4], namely $Z = \Pi(Z, E)$.

The operator $\mathcal{L}$ is a special case of a Sylvester operator. It is singular if and only if the matrices $A - X_0$ and $A + X_0$ have a common eigenvalue [5].

The eigenvalues of $\mathcal{L}$ are the eigenvalues of its matrix L and they are equal to $\lambda_i(A - X_0) - \lambda_j(A + X_0), i, j = 1, 2, ..., n$, where $\lambda_1(H)$, $\lambda_2(H)$, …, $\lambda_n(H)$ are the eigenvalues of the matrix $H \in \mathbb{R}^{n\times n}$ counted according to their algebraic multiplicities.

Hence the operator $\mathcal{L}$ and its matrix L would be invertible if and only if $\lambda_i(A - X_0) \neq \lambda_j(A + X_0), i, j = 1, 2, ..., n$. However, for equation (3) with a matrix A having multiple eigenvalues and a solution X_0 being a nilpotent matrix, the operator $\mathcal{L}$, as defined by (5), is singular. Hence equation (4) may not be written immediately as an equivalent operator equation. As a consequence, the standard technique for perturbation analysis of matrix equations [3, 4] may not be implemented.

Rewrite the matrix equation (4) in a vector form applying the vec operation to the first order terms O(u) and having in mind that $F(X_0, A) = 0$:

$$L\mathrm{vec}(Z) = -M\mathrm{vec}(E) + \mathrm{O}(u^2).$$

Let the rank of the matrix L be $r \geq 1$ and consider the singular value decomposition $L = U \Sigma V^{\mathrm{T}}$ of L, where U and V are $n^2 \times n^2$ orthogonal matrices, $\Sigma = \mathrm{diag}(\Sigma_1, 0)$, $\Sigma_1 = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$, and $\sigma_1 \geq \sigma_2 \geq \ldots, \sigma_r > 0$ are the positive singular values of L.

Denote

$$P_1 := \begin{bmatrix} I_r\, 0 \end{bmatrix} \in \mathbb{R}^{r \times n^2},\ P_2 := \begin{bmatrix} 0 I_{n^2 - r} \end{bmatrix} \in \mathbb{R}^{(n^2 - r) \times n^2},$$
$$\Pi_1 := \begin{bmatrix} P_1 \\ 0 \end{bmatrix} \in \mathbb{R}^{n^2 \times n^2},\ \Pi_2 := \begin{bmatrix} 0 \\ P_2 \end{bmatrix} \in \mathbb{R}^{n^2 \times n^2}$$

and

$$z := V^{\mathrm{T}} \mathrm{vec}(Z) = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \in \mathbb{R}^{n^2},\ z_k := P_k z,$$
$$e := -U^{\mathrm{T}} M \mathrm{vec}(E) = \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \in \mathbb{R}^{n^2},\ e_k := -P_k U^{\mathrm{T}} M \mathrm{vec}(E),\ k = 1, 2.$$

Then we obtain

$$\Sigma_1 z_1 = e_1 + O(u^2),$$
$$0 = e_2 + O(u^2).$$

Hence

$$z_1 = \Sigma_1^{-1} e_1 + O(u^2)$$

and

$$\|z_1\|_2 = \|\Sigma_1^{-1} e_1\|_2 + O(u^2) = \|\Sigma_1^{-1} P_1 U^{\mathrm{T}} M \mathrm{vec}(E)\|_2 + O(u^2) \leq$$
$$\leq \|\Sigma_1^{-1} P_1 U^{\mathrm{T}} M\|_2 \|\mathrm{vec}(E)\|_2 + O(u^2).$$

Hence we have derived the following first order bound for the norm of the projection $\Pi_1 V^{\mathrm{T}} \mathrm{vec}(Z)$ of the vectorization $\mathrm{vec}(Z)$ of the perturbation Z in the solution X_0

(7)
$$\|\Pi_1 V^{\mathrm{T}} \mathrm{vec}(Z)\|_2 = \|P_1 V^{\mathrm{T}} \mathrm{vec}(Z)\|_2 \leq C\varepsilon,$$
$$C = C(A, X_0) := \|\Sigma_1^{-1} P_1 U^{\mathrm{T}} M\|_2 \leq \frac{\|M\|_2}{\sigma_r},$$

where $\varepsilon = \|E\|_F = \|\mathrm{vec}(E)\|_2$.

The local bound (7) is valid only asymptotically, for $\varepsilon \to 0$. This means that the perturbation in the data must be small enough to ensure sufficient accuracy of the local bound. Unfortunately, it is usually impossible to say, having a small but finite perturbation ε, whether the neglected terms are indeed negligible.

The disadvantages of the local bound may be overcome using the techniques of non-local perturbation analysis.

4. Non-local perturbation analysis

Equation (3) may be written in the form

(8)
$$\mathcal{L}(Z) = -\mathcal{M}(E) + E Z - Z E + Z^2.$$

The vector representation of this equation is

$$L\, \mathrm{vec}(Z) = -M \mathrm{vec}(E) + \mathrm{vec}(EZ - ZE + Z^2).$$

Using the notations from the previous section we may rewrite the last equation as

$$z_1 = \Phi_1(z, E) := \Sigma_1^{-1} e_1 + \Sigma_1^{-1} P_1 U^{\mathrm{T}} \mathrm{vec}(EZ - ZE + Z^2),$$
$$0 = e_2 + P_2 U^{\mathrm{T}} \mathrm{vec}(EZ - ZE + Z^2),$$

where

$$z = V^{\mathrm{T}} \mathrm{vec}(Z) = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, z_k = P_k z, Z = \mathrm{vec}^{-1}(Vz).$$

Setting $\Phi_2(z) := z_2$ we see that z satisfies the operator equation

$$z = \Phi(z, E) := \begin{bmatrix} \Phi_1(z, E) \\ \Phi_2(z) \end{bmatrix}.$$

For fixed numbers $\rho > 0$ and $\nu \in (0, 1]$ let $\mathcal{V}_\nu(\rho) \subset \mathbb{R}^{n^2}$ be the set of vectors z such that $\|z\|_2 \le \rho$ and $\|z_1\|_2 = \|\Pi_1 z\|_2 \le \nu\rho$. This set is closed and convex.

Next we shall find conditions on the norm $\varepsilon = \|E\|_F$ which guarantee the existence of a quantity $\rho_0 > 0$ such that $\Phi(\mathcal{V}_\nu(\rho_0), E) \subset \mathcal{V}_\nu(\rho_0)$. For $z \in \mathcal{V}_\nu(\rho)$ we have

$$\|\Phi_1(z, E)\|_2 \le h_\nu(\rho, \varepsilon) := \frac{\rho^2}{\sigma_r} + \frac{2\varepsilon\rho}{\sigma_r} + C\varepsilon .$$

Suppose that

(9) $$\varepsilon \le \varepsilon_\nu := \frac{\sigma_r \nu^2}{(\sqrt{C} + \sqrt{C + 2\nu})^2} .$$

Then we may define the quantity

(10) $$\rho_0 = f_\nu(\varepsilon) := \frac{2\sigma_r C\varepsilon}{\sigma_r \nu - 2\varepsilon + \sqrt{(\sigma_r \nu - 2\varepsilon)^2 - 4\sigma_r C\varepsilon}} .$$

For $z \in \mathcal{V}_\nu(\rho_0)$ we shall have $h_\nu(\rho_0, \varepsilon) = \nu\rho_0$ and hence the operator $\Phi(.,E)$ transforms the set $\mathcal{V}_\nu(\rho_0)$ into itself. Then, according to the Schauder fixed point principle, the operator $\Phi(.,E)$ has a fixed point $z \in \mathcal{V}_\nu(\rho_0)$ for which the estimate

$$\|z_1\|_2 \le \nu f_\nu(\varepsilon), \varepsilon \in [0, \varepsilon_\nu],$$

holds. Moreover, in this case we have the following result.

Theorem 3. Let the quantity $\nu \in (0, \varepsilon_\nu]$ be given and let $\varepsilon \in [0, \varepsilon_\nu]$. Then there exists a perturbation Z in X_0 such that

(11) $$\|Z\|_F = \|z\|_2 \le f_\nu(\varepsilon),$$

where ε_ν and $f_\nu(\varepsilon)$ are determined by relations (9), (10) and (7).

5. Numerical examples

In this section we give three numerical examples to illustrate the results from Sections 3 and 4.

Example 4. Consider the matrix equation $XA - AX = X^2$ from Example 2.8 in [1] with a data matrix A and a solution X_0 given by

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, X_0 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

The perturbations E in the data and Z in the solution are taken as

$$E = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & s \\ 0 & 0 & 0 \end{bmatrix}, Z = \begin{bmatrix} 0 & 0 & 0 \\ -s & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \text{ with } s = 10^{-2k} \text{ for } k = 1, 2, 3, 4, 5.$$

We estimate the Euclidean norm of the projection $\Pi_1 V^{\mathrm{T}}\mathrm{vec}(Z)$ of the perturbation Z in the solution X_0 by the local bound $C\varepsilon$ from (7). Then we estimate this quantity by the bound $\nu f_\nu(\varepsilon)$ using (11). The results obtained for different values of k and ν are shown in Table 1.

Table 1. Perturbation bounds for $z_1 = P_1 V^{\mathrm{T}}$ vec(Z) (Example 4)

k	$\lVert z_1 \rVert_2$	$C\varepsilon$	$\nu f_\nu(\varepsilon)$, $\nu = 0.25$	$\nu f_\nu(\varepsilon)$, $\nu = 0.5$	$\nu f_\nu(\varepsilon)$, $\nu = 0.75$	$\nu f_\nu(\varepsilon)$, $\nu = 1$
1	6.3246×10^{-3}	3.6187×10^{-2}	*	5.0938×10^{-2}	4.1046×10^{-2}	3.8910×10^{-2}
2	6.3246×10^{-5}	3.6187×10^{-4}	3.6474×10^{-4}	3.6267×10^{-4}	3.6226×10^{-4}	3.6211×10^{-4}
3	6.3246×10^{-7}	3.6187×10^{-6}	3.6190×10^{-6}	3.6188×10^{-6}	3.6187×10^{-6}	3.6187×10^{-6}
4	6.3246×10^{-9}	3.6187×10^{-8}	3.6187×10^{-8}	3.6187×10^{-8}	3.6187×10^{-8}	3.6187×10^{-8}
5	6.3246×10^{-11}	3.6187×10^{-10}	3.6187×10^{-10}	3.6187×10^{-10}	3.6187×10^{-10}	3.6187×10^{-10}

Next we estimate the Frobenius norm of Z by the non-local bound $f_\nu(\varepsilon)$ from (11). The results are given in Table 2.

The cases when the non-local bound is not valid, since the existence condition (9) is violated, are denoted by asterisks.

Table 2. Perturbation bounds for Z (Example 4)

k	$\lVert Z \rVert_F$	$f_\nu(\varepsilon)$, $\nu = 0.25$	$f_\nu(\varepsilon)$, $\nu = 0.5$	$f_\nu(\varepsilon)$, $\nu = 0.75$	$f_\nu(\varepsilon)$, $\nu = 1$
1	1.0000×10^{-2}	*	1.0188×10^{-1}	5.4727×10^{-2}	3.8910×10^{-2}
2	1.0000×10^{-4}	1.4590×10^{-3}	7.2533×10^{-4}	4.8301×10^{-4}	3.6211×10^{-4}
3	1.0000×10^{-6}	1.4476×10^{-5}	7.2375×10^{-6}	4.8250×10^{-6}	3.6187×10^{-6}
4	1.0000×10^{-8}	1.4475×10^{-7}	7.2374×10^{-8}	4.8249×10^{-8}	3.6187×10^{-8}
5	1.0000×10^{-10}	1.4475×10^{-9}	7.2374×10^{-10}	4.8249×10^{-10}	3.6187×10^{-10}

As it is seen the non-local bound $\nu f_\nu(\varepsilon)$ is slightly more pessimistic than the local bound $C\varepsilon$.

Example 5. Consider the matrix equation (1) with matrices

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & a \\ 0 & 0 & 0 \end{bmatrix}, \quad X_0 = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1-a \\ -1 & 0 & 1 \end{bmatrix} \text{ with } a = \frac{1}{3} - 10^{-k}.$$

The perturbations E and Z are taken as in the previous example.

The estimated Euclidean norm of the projection $\Pi_1 V^{\mathrm{T}}\mathrm{vec}(Z)$ of the perturbation Z in the solution X_0, the local bound $C\varepsilon$, defined by (7), and the non-local bound $\nu f_\nu(\varepsilon)$ from (11) for k=1, 2, 3, 4, 5 and ν = 0.25, 0.5, 0.75, 1 are shown in Table 3.

Table 3. Perturbation bounds for $z_1 = P_1 V^T$ vec(Z) (Example 5)

k	$\|\|z_1\|\|_2$	$C\varepsilon$	$\nu f_\nu(\varepsilon)$, ν = 0.25	$\nu f_\nu(\varepsilon)$, ν = 0.5	$\nu f_\nu(\varepsilon)$, ν = 0.75	$\nu f_\nu(\varepsilon)$, ν = 1
1	6.3246×10^{-3}	1.7923×10^{-1}	*	*	*	*
2	6.3246×10^{-5}	1.2500×10^{-3}	1.3678×10^{-3}	1.2767×10^{-3}	1.2621×10^{-3}	1.2570×10^{-3}
3	6.3246×10^{-7}	1.2123×10^{-5}	1.2132×10^{-5}	1.2125×10^{-5}	1.2124×10^{-5}	1.2124×10^{-5}
4	6.3246×10^{-9}	1.2086×10^{-7}	1.2087×10^{-7}	1.2086×10^{-7}	1.2086×10^{-7}	1.2086×10^{-7}
5	6.3246×10^{-11}	1.2083×10^{-9}	1.2083×10^{-9}	1.2083×10^{-9}	1.2083×10^{-9}	1.2083×10^{-9}

The results of the estimation of the Frobenius norm of Z by the non-local bound $f_\nu(\varepsilon)$ from (11) for different values of k and ν are shown in Table 4.

Table 4. Perturbation bounds for Z (Example 5)

k	$\|\|Z\|\|_F$	$f_\nu(\varepsilon)$, ν = 0.25	$f_\nu(\varepsilon)$, ν = 0.5	$f_\nu(\varepsilon)$, ν = 0.75	$f_\nu(\varepsilon)$, ν = 1
1	1.0000×10^{-2}	*	*	*	*
2	1.0000×10^{-4}	5.4713×10^{-3}	2.5534×10^{-3}	1.6827×10^{-3}	1.2570×10^{-3}
3	1.0000×10^{-6}	4.8528×10^{-5}	2.4251×10^{-5}	1.6165×10^{-5}	1.2124×10^{-5}
4	1.0000×10^{-8}	4.8346×10^{-7}	2.4173×10^{-7}	1.6115×10^{-7}	1.2086×10^{-7}
5	1.0000×10^{-10}	4.8331×10^{-9}	2.4166×10^{-9}	1.6110×10^{-9}	1.2083×10^{-9}

The cases when the non-local bound is not valid, since the existence condition (9) does not hold, are denoted by asterisks.

Example 6. Consider the matrix equation (1) with matrices

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & a \\ 0 & 0 & 1 \end{bmatrix}, X_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & a \\ 0 & 0 & 0 \end{bmatrix} \text{ with } a = \frac{1}{3} - 10^{-k}.$$

Suppose that the perturbation E is

$$E = \begin{bmatrix} s & 0 & 0 \\ 0 & s & s \\ 0 & 0 & s \end{bmatrix} \text{ with } s = 10^{-2k}.$$

The perturbation Z in the solution X_0 is the same as in Examples 4 and 5.

The results obtained for k = 1, 2, 3, 4, 5 and for ν = 0.25, 0.5, 0.75, 1 for the estimated quantity $\|z_1\|_2 = \| \Pi_1 V^T \text{vec}(Z) \|_2$ and for the bounds $C\varepsilon$ from (7), and $f_\nu(\varepsilon)$ from (11), are shown in Table 5.

Table 5. Perturbation bounds for $z_1 = P_1 V^T$ vec(Z) (Example 6)

k	$\|\|z_1\|\|_2$	$C\varepsilon$	$\nu f_\nu(\varepsilon)$, ν = 0.25	$\nu f_\nu(\varepsilon)$, ν = 0.5	$\nu f_\nu(\varepsilon)$, ν = 0.75	$\nu f_\nu(\varepsilon)$, ν = 1
1	0	1.4142×10^{-2}	*	2.2138×10^{-2}	1.7246×10^{-2}	1.6073×10^{-2}
2	0	1.4142×10^{-4}	1.4227×10^{-4}	1.4172×10^{-4}	1.4159×10^{-4}	1.4154×10^{-4}
3	0	1.4142×10^{-6}	1.4143×10^{-6}	1.4142×10^{-6}	1.4142×10^{-6}	1.4142×10^{-6}
4	0	1.4142×10^{-8}	1.4142×10^{-8}	1.4142×10^{-8}	1.4142×10^{-8}	1.4142×10^{-8}
5	0	1.4142×10^{-10}	1.4142×10^{-10}	1.4142×10^{-10}	1.4142×10^{-10}	1.4142×10^{-10}

The results for the non-local perturbation bound $f_\nu(\varepsilon)$ (11) for the norm of Z are given in Table 6.

Table 6. Perturbation bounds for Z (Example 6)

k	$\|\|Z\|\|_F$	$f_\nu(\varepsilon)$, $\nu = 0.25$	$f_\nu(\varepsilon)$, $\nu = 0.5$	$f_\nu(\varepsilon)$, $\nu = 0.75$	$f_\nu(\varepsilon)$, $\nu = 1$
1	1.0000×10^{-2}	*	4.4276×10^{-2}	2.2995×10^{-2}	1.6073×10^{-2}
2	1.0000×10^{-4}	5.6910×10^{-4}	2.8344×10^{-4}	1.8879×10^{-4}	1.4154×10^{-4}
3	1.0000×10^{-6}	5.6572×10^{-6}	2.8285×10^{-6}	1.8856×10^{-6}	1.4142×10^{-6}
4	1.0000×10^{-8}	5.6569×10^{-8}	2.8284×10^{-8}	1.8856×10^{-8}	1.4142×10^{-8}
5	1.0000×10^{-10}	5.6569×10^{-10}	2.8284×10^{-10}	1.8856×10^{-10}	1.4142×10^{-10}

As it is seen, here the local bound $C\varepsilon$ estimates a projection of the perturbation Z in the solution X_0, which in this particular example is the zero vector.

6. Concluding remarks

In this paper a perturbation analysis of the matrix equation $XA - AX = X^2$ is presented. Local and non-local perturbation bounds are derived under the Assumptions A2-A4, fulfilled for a set of small perturbations preserving the Jordan form of A. The local bound concerns only a projection of the perturbation in the solution and gives satisfactory results for small perturbations in the data. The non-local bound is slightly more pessimistic but holds when the perturbation in the data belongs to a preliminary defined domain of applicability of the bound. Numerical examples demonstrate the effectiveness of the bounds proposed.

References

1. B u r d e, K. On the matrix equation $XA - AX = X^p$. – Linear Algebra Appl., Vol. **404**, 2005, 147-165.
2. I k r a m o v, K. D. A Note on the Matrix Equation $XA - AX = X^p$. – Linear Algebra Appl., Vol. **427**, 2007, 390-392.
3. K o n s t a n t i n o v, K., D. W. G u, V. M e h r m a n n, P. P e t k o v. Perturbation Theory for Matrix Equations. ISBN 0-444-51315-9, North-Holland, Amsterdam, 2003.
4. K o n s t a n t i n o v, M., V. M e h r m a n n, P. P e t k o v, D. W. G u. A General Framework for the Perturbation Theory of Matrix Equations. Technical Report 760, Inst. Math., TU – Berlin, 2002.
http://www.math.tu-berlin.de/preprints
5. K o n s t a n t i n o v, M., M. S t a n i s l a v o v a, P. P e t k o v. Perturbation Bounds and Characterisation of the Solution of the Associated Algebraic Riccati Equation. – Linear Algebra Appl., Vol. **285**, 1998, 7-31.
6. L a n c a s t e r, P., L. R o d m a n. The Algebraic Riccati Equation. Oxford, Oxford Univ. Press, 1995.