East Asian Mathematical Journal
Vol. 29 (2013), No. 3, pp. 327–335
http://dx.doi.org/10.7858/eamj.2013.023

THE CONDITION NUMBERS OF A QUADRATIC MATRIX EQUATION*

Hye-Yeon Kim† and Hyun-Min Kim‡

Abstract. In this paper we consider the quadratic matrix equation which can be defined by

$$Q(X) = AX^2 + BX + C = 0,$$

where X is a $n \times n$ unknown complex matrix, and A, B and C are $n \times n$ given matrices with complex elements. We first introduce a couple of condition numbers of the equation $Q(X)$ and present normwise condition numbers. Finally, we compare the results and some numerical experiments are given.

1. Introduction

In this paper, we consider some different typical condition numbers of the quadratic matrix equation:

$$Q(X) = AX^2 + BX + C = 0 \tag{1}$$

where $A, B, C \in \mathbb{C}^{n\times n}$. If a matrix S satisfies the equation $Q(S) = 0$, then S is called a solvent of $Q(X)$. The condition number is important in numerical sense because it provides information about sensitivity of the solution to perturbations in the data. The theories for finding the condition number for (1) were suggested by Davis [1]. He considered traditional condition numbers which derived and expressed using norms. Also, the mixed perturbation analysis was suggested by Skeel [6] and he obtained the mixed analysis for Gaussian elimination. Higham and Kim [3] considered the componentwise perturbation theory for the equation $Q(X)$. Gohberg and Koltracht [2] obtained explicit expressions for both mixed and componentwise condition numbers. For the

Received April 4, 2013; Accepted April 18, 2013.
2000 *Mathematics Subject Classification.* 58B34, 58J42, 81T75.
Key words and phrases. quadratic matrix equation, condition number, backward analysis.
*This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology(2012R1A1A2008840).
†This work is based on the first author's M.Sc. thesis.
‡Corresponding author.

generalized Sylvester equation, Lin and Wei [5] gave the mixed and componentwise condition numbers. In this work, we consider normwise condition numbers for a solvent to the quadratic matrix equation (1).

2. Classical condition numbers

First, the Fréchet derivative of the quadratic matrix equation $Q(X)$ in (1) can be derived by

$$\begin{aligned} Q(X+H) &= A(X+H)^2 + B(X+H) + C \\ &= Q(X) + ((AX+B)H + AHX) + AH^2 \\ &= Q(X) + Q'(X)H + AH^2. \end{aligned}$$

By applying the Kronecker product $A \otimes B = (a_{ij}B)$ and the vec operator property [4] to $Q'(X)H$, we can obtain

$$\begin{aligned} \mathrm{vec}(Q'(X)H) &= (I_n^T \otimes (AX+B))\mathrm{vec}(H) + (X^T \otimes A)\mathrm{vec}(H) \\ &= [(I_n^T \otimes (AX+B)) + (X^T \otimes A)]\mathrm{vec}(H). \end{aligned}$$

We now introduce two results for finding the condition numbers of a solvent to quadratic matrix equation $Q(X)$ in (1).

Assumption.

(i) *Let X be an exact solvent of $Q(X) = AX^2 + BX + C = 0$,*

(ii) *Let $\hat{X} = X + \Delta X$ be a solvent of the perturbed equation*

$$\hat{Q}(\hat{X}) = \hat{A}\hat{X}^2 + \hat{B}\hat{X} + \hat{C} = 0,$$

(iii) $\hat{A} = A + \Delta A,\ \hat{B} = B + \Delta B,\ \hat{C} = C + \Delta C$ *with* $\|\Delta A\|_F \le \varepsilon\|A\|_F$, $\|\Delta B\|_F \le \varepsilon\|B\|_F, \|\Delta C\|_F \le \varepsilon\|C\|_F$,

(iv) $\|Q'(X)^{-1}\|_F \cdot \|\Delta Q'(X)\|_F \le k < 1$, *where* $\Delta Q'(X) = \hat{Q}'(X) - Q'(X)$

Under the Assumption, we can get the following results on the error in X.

Theorem 2.1. [1] *For sufficiently small ε, $\|\Delta X\|_F \le \gamma\varepsilon$, where*

$$\begin{aligned} \alpha &= \frac{2}{1-k}\|Q'(X)^{-1}\|_F[\|A\|_F \cdot \|X\|_F^2 + \|B\|_F \cdot \|X\|_F], \\ \beta &= \frac{1+\varepsilon}{1-k}\|Q'(X)^{-1}\|_F \cdot \|A\|_F, \\ \gamma &= \frac{2\alpha}{1+\sqrt{1-4\alpha\beta\varepsilon}} = \alpha + O(\varepsilon^2). \end{aligned}$$

Theorem 2.2. [3]

$$\frac{\|\Delta X\|_F}{\|X\|_F} \le \Psi(X)\varepsilon + O(\varepsilon^2),$$

where

$$\begin{aligned} \varepsilon &= \|[\alpha^{-1}\Delta A,\ \beta^{-1}\Delta B,\ \gamma^{-1}\Delta C]\|_F, \\ \Psi(X) &= \|P^{-1}[\alpha(X^2)^T \otimes I_n,\ \beta X^T \otimes I_n,\ \gamma I_{n^2}]\|_2 / \|X\|_F, \\ P &= I_n^T \otimes (AX+B) + X^T \otimes A. \end{aligned}$$

and ΔX is perturbation of X due to the perturbations $\Delta A, \Delta B, \Delta C$ of A, B, C.

3. Normwise condition numbers

We now consider normwise condition numbers. Let the mapping

$$\varphi : (A, B, C) \mapsto \mathrm{vec}(X)$$

where X is the solvent of equation $Q(X)$ in (1). The normwise condition numbers of the quadratic matrix equation can be defined by

$$\kappa_1(\varphi) = \lim_{\varepsilon\to 0} \sup_{\Delta_1 \le \varepsilon} \frac{\|\Delta X\|_F}{\varepsilon \|X\|_F},$$

$$\kappa_2(\varphi) = \lim_{\varepsilon\to 0} \sup_{\Delta_2 \le \varepsilon} \frac{\|\Delta X\|_F}{\varepsilon \|X\|_F}$$

and

$$\kappa_3(\varphi) = \lim_{\varepsilon\to 0} \sup_{\Delta_3 \le \varepsilon} \frac{\|\Delta X\|_F}{\varepsilon \|X\|_F},$$

where ΔX is the perturbation of X due to the perturbations $\Delta A, \Delta B$ and ΔC of A, B and C, and

$$\begin{aligned}
\Delta_1 &= \left\| \left[\frac{\|\Delta A\|_F}{\alpha}, \frac{\|\Delta B\|_F}{\beta}, \frac{\|\Delta C\|_F}{\gamma} \right] \right\|_2, \\
\Delta_2 &= \max\left\{ \frac{\|\Delta A\|_F}{\alpha}, \frac{\|\Delta B\|_F}{\beta}, \frac{\|\Delta C\|_F}{\gamma} \right\}, \\
\Delta_3 &= \frac{\|[\|\Delta A\|_F, \|\Delta B\|_F, \|\Delta C\|_F]\|_2}{\|[\alpha, \beta, \gamma]\|_2},
\end{aligned}$$

for $\alpha = \|A\|_F, \beta = \|B\|_F, \gamma = \|C\|_F$.

The next result shows upper bounds of each condition number.

Theorem 3.1. *By the above notation, the normwise condition number of the quadratic matrix equation are*

$$\begin{aligned}
&\text{(i) } \kappa_1(\varphi) &&\le \frac{\|P^{-1}S_1\|_2}{\|X\|_F}, \\
&\text{(ii) } \kappa_2(\varphi) &&\le \min\{\kappa^U(\varphi), \kappa^M(\varphi)\}, \\
&\text{(iii) } \kappa_3(\varphi) &&\le \frac{\|P^{-1}S_2\|_2}{\|X\|_F} \left(\|A\|_F^2 + \|B\|_F^2 + \|C\|_F^2\right)^{\frac{1}{2}},
\end{aligned}$$

where

$$\begin{aligned}
P &= I_n^T \otimes AX + X^T \otimes A + I_n^T \otimes B, \\
S_1 &= \left[\alpha (X^2)^T \otimes I_n, \quad \beta X^T \otimes I_n, \quad \gamma I_{n^2}\right], \\
S_2 &= \left[(X^2)^T \otimes I_n, \quad X^T \otimes I_n, \quad I_{n^2}\right],
\end{aligned}$$

$$\kappa^U(\varphi) = \frac{u}{\|X\|_F}, \quad \kappa^M(\varphi) = \frac{m}{\|X\|_F},$$

and

$$\begin{aligned} u &= \sqrt{3}\|P^{-1}S_1\|_2, \\ m &= \alpha\|P^{-1}[(X^2)^T \otimes I_n]\|_2 + \beta\|P^{-1}[X^T \otimes I_n]\|_2 + \gamma\|P^{-1}\|_2. \end{aligned}$$

Proof. The perturbed equation of the quadratic matrix equation in (1) is

$$(A + \Delta A)(X + \Delta X)^2 + (B + \Delta B)(X + \Delta X) + C + \Delta C = 0. \tag{2}$$

Neglecting second-order terms in (2) we obtain the equation

$$AX\Delta X + A\Delta XX + B\Delta X = -\Delta AX^2 - \Delta BX - \Delta C.$$

Applying the vec operator to both sides the equation can be written by

$$\begin{aligned} &(I_n^T \otimes AX + X^T \otimes A + I_n^T \otimes B)\mathrm{vec}(\Delta X) \\ &= -((X^2)^T \otimes I_n)\mathrm{vec}(\Delta A) - (X^T \otimes I_n)\mathrm{vec}(\Delta B) - (I_n^T \otimes I_n)\mathrm{vec}(\Delta C). \end{aligned}$$

Let $P = I_n^T \otimes AX + X^T \otimes A + I_n^T \otimes B$ then

$$\begin{aligned} P\mathrm{vec}(\Delta X) &= -\begin{bmatrix}\alpha(X^2)^T \otimes I_n, & \beta_2 X^T \otimes I_n, & \gamma_3 I_{n^2}\end{bmatrix}\begin{bmatrix}\mathrm{vec}(\Delta A)/\alpha \\ \mathrm{vec}(\Delta B)/\beta \\ \mathrm{vec}(\Delta C)/\gamma\end{bmatrix} \\ &= -S_1 r, \end{aligned}$$

where

$$r = \begin{bmatrix}\mathrm{vec}(\Delta A)/\alpha \\ \mathrm{vec}(\Delta B)/\beta \\ \mathrm{vec}(\Delta C)/\gamma\end{bmatrix}.$$

Finally, we have the equation

$$\mathrm{vec}(\Delta X) = -P^{-1}S_1 r. \tag{3}$$

By taking 2-norm we obtain

$$\|\mathrm{vec}(\Delta X)\|_2 = \|\Delta X\|_F \le \|P^{-1}S_1\|_2\|r\|_2. \tag{4}$$

If $\|r\|_2 \le \varepsilon$ then

$$\begin{aligned} \kappa_1(\varphi) &= \lim_{\varepsilon\to 0}\sup_{\Delta_1\le\varepsilon}\frac{\|\Delta X\|_F}{\varepsilon\|X\|_F} \\ &\le \lim_{\varepsilon\to 0}\sup_{\Delta_1\le\varepsilon}\frac{\|P^{-1}S_1\|_2\|r\|_2}{\varepsilon\|X\|_F} \\ &\le \lim_{\varepsilon\to 0}\sup_{\Delta_1\le\varepsilon}\frac{\|P^{-1}S_1\|_2\varepsilon}{\varepsilon\|X\|_F} \\ &= \frac{\|P^{-1}S_1\|_2}{\|X\|_F}. \end{aligned}$$

Now, let $\varepsilon = \max\left\{\dfrac{\|\Delta A\|_F}{\alpha}, \dfrac{\|\Delta B\|_F}{\beta}, \dfrac{\|\Delta C\|_F}{\gamma}\right\}$, then

$$\begin{aligned}\|\Delta X\|_F &\le \|P^{-1}S_1\|_2\|r\|_2 \\ &= \|P^{-1}S_1\|_2\left[\frac{\|\Delta A\|_F^2}{\alpha^2} + \frac{\|\Delta B\|_F^2}{\beta^2} + \frac{\|\Delta C\|_F^2}{\gamma^2}\right]^{\frac{1}{2}} \\ &\le \sqrt{3}\varepsilon\|P^{-1}S_1\|_2 = \varepsilon u.\end{aligned} \tag{5}$$

And from (3),

$$\begin{aligned}\mathrm{vec}(\Delta X) &= -\alpha P^{-1}[(X^2)^T \otimes I_n]\frac{\mathrm{vec}(\Delta A)}{\alpha} - \beta P^{-1}[X^T \otimes I_n]\frac{\mathrm{vec}\Delta B}{\beta} \\ &\quad -\gamma P^{-1}I_{n^2}\frac{\mathrm{vec}(\Delta C)}{\gamma}.\end{aligned}$$

Using $\|\mathrm{vec}(A)\|_2 = \|A\|_F$,

$$\begin{aligned}\|\Delta X\|_F &\le \alpha\|P^{-1}[(X^2)^T \otimes I_n]\|_2\left\|\frac{\mathrm{vec}(\Delta A)}{\alpha}\right\|_2 \\ &\quad +\beta\|P^{-1}[X^T \otimes I_n]\|_2\left\|\frac{\mathrm{vec}(\Delta B)}{\beta}\right\|_2 \\ &\quad +\delta_3\|P^{-1}I_{n^2}\|_2\left\|\frac{\mathrm{vec}(\Delta C)}{\gamma}\right\|_2 \\ &\le \alpha\|P^{-1}[(X^2)^T \otimes I_n]\|_2\left[\frac{\|\Delta A\|_F^2}{\alpha^2}\right]^{\frac{1}{2}} \\ &\quad +\beta\|P^{-1}[X^T \otimes I_n]\|_2\left[\frac{\|\Delta B\|_F^2}{\beta^2}\right]^{\frac{1}{2}} + \gamma\|P^{-1}\|_2\left[\frac{\|\Delta C\|_F^2}{\gamma^2}\right]^{\frac{1}{2}} \\ &\le \left(\alpha\|P^{-1}[(X^2)^T \otimes I_n]\|_2 + \beta\|P^{-1}[X^T \otimes I_n]\|_2 + \beta\|P^{-1}\|_2\right) \\ &\quad \max\left\{\frac{\|\Delta A\|_F}{\alpha}, \frac{\|\Delta B\|_F}{\beta}, \frac{\|\Delta C\|_F}{\delta_3}\right\} \\ &= m\varepsilon.\end{aligned} \tag{6}$$

Then, (5) and (6) imply Theorem 2.3 (ii). Let $\varepsilon = \dfrac{\|[\|\Delta A\|_F, \|\Delta B\|_F, \|\Delta C\|_F]\|_2}{\|[\|A\|_F, \|B\|_F, \|C\|_F]\|_2}$ then

$$\begin{aligned}\kappa_3(\varphi) &= \lim_{\varepsilon\to 0}\sup_{\Delta_3\le\varepsilon}\frac{\|\Delta X\|_F}{\varepsilon\|X\|_F} \\ &\le \lim_{\varepsilon\to 0}\sup_{\Delta_3\le\varepsilon}\frac{\|P^{-1}S_2\|_2\|r\|_2}{\varepsilon\|X\|_F} \\ &= \lim_{\varepsilon\to 0}\sup_{\Delta_3\le\varepsilon}\frac{\|P^{-1}S_2\|_2[\|A\|_F^2 + \|B\|_F^2 + \|C\|_F^2]^{\frac{1}{2}}}{\varepsilon\|X\|_F} \\ &= \frac{\|P^{-1}S_2\|_2}{\|X\|_F}[\|A\|_F^2 + \|B\|_F^2 + \|C\|_F^2]^{\frac{1}{2}}.\end{aligned}$$

□

If we choose $\alpha = \|A\|_F, \beta = \|B\|_F, \gamma = \|C\|_F$, then we have $\kappa_1(\varphi) \leq \kappa_2(\varphi) \leq \kappa_3(\varphi)$ by (3) and $\kappa_1(\varphi)$ is same to the result of Higham and Kim [3].

4. Numerical Experiments

In this section, we show and compare numerical experiments by applying our results. All experiments are done in MATLAB 7.1 and all iterations terminated when the relative residual $\rho_Q(X_k)$ satisfies

$$\rho_Q(X_k) = \frac{\|fl(Q(X_i))\|_F}{\|A\|_F\|X_k\|_F^2 + \|B\|_F\|X_k\|_F + \|C\|_F} \leq n\varepsilon,$$

where $\varepsilon = 1e-016$.

First, we consider a 2×2 quadratic matrix equation with well-conditioned.

Example 4.1. *The quadratic matrix equation is given by*

$$\begin{aligned} Q_1(X) &= A_1X^2 + B_1X + C_1 \\ &= \begin{bmatrix} 0.002 & 0 \\ 0 & 0.002 \end{bmatrix} X^2 + \begin{bmatrix} -0.002 & -0.006 \\ -0.006 & -0.002 \end{bmatrix} X + \begin{bmatrix} 0 & 0.006 \\ 0.006 & 0 \end{bmatrix} = 0. \end{aligned}$$

Suppose that the equation perturbed with $\Delta A_1 = \begin{bmatrix} 0.00050 & 0 \\ 0 & 0 \end{bmatrix}$, $\Delta B_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0.00050 \end{bmatrix}$, $\Delta C_1 = \begin{bmatrix} -0.41205 & 0 \\ 0.24 & 0.00050 \end{bmatrix}$ respectively. Then using the Newton's method, we obtain the solvent matrix $\hat{X} = \begin{bmatrix} 41 & 0 \\ 0 & 1 \end{bmatrix}$ of the perturbed equation

$$\hat{Q}_1(\hat{X}) = (A_1 + \Delta A_1)\hat{X}^2 + (B_1 + \Delta B_1)\hat{X} + (C_1 + \Delta C_1) = 0, \tag{7}$$

where $\hat{X} = X + \Delta X$. Now we can get the condition number of (7). When we use the result of Theorem 2.1, the upper bound of condition is 13274. Also 729.6705 is provided by Theorem 2.2. This example shows that the result of Theorem 2.1 is sharper than Theorem 2.1. In this case, starting Newton's method with $X_0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, we have the results displayed in Table 1. Then we can get the solvent matrix $\hat{X} = \begin{bmatrix} 41 & 0 \\ 0 & 1 \end{bmatrix}$ of (7) and since the perturbed quadratic matrix equation $\hat{Q}_1(\hat{X}) = 10^{-18} \begin{bmatrix} 0 & 0 \\ 0 & 0.2168 \end{bmatrix}$, $\hat{X}$ is exact solvent of $\hat{Q}_1(\hat{X})$.

No.iteration	$\rho_Q(X_k)$ of Newton's Method
1	3.96e-002
2	6.42e-003
3	1.04e-003
4	1.58e-004
5	1.63e-005
6	3.40e-007
7	1.85e-010
8	1.10e-016

Table 1. Relative residual for problem (7) with Newton's method

We now consider an example with 3×3 coefficient matrices. Through this example, we see the necessity of the condition number.

Example 4.2. *The quadratic matrix equation is given by*

$$\begin{aligned} Q_2(X) &= \begin{bmatrix} 0.0430 & 0.7803 & 0.3667 \\ 0.3820 & 0.4279 & 0.9778 \\ 0.6368 & 0.1712 & 0.4593 \end{bmatrix} X^2 + \begin{bmatrix} 0.8541 & 0.9208 & 0.3445 \\ 0.1521 & 0.9804 & 0.2094 \\ 0.4425 & 0.4635 & 0.3162 \end{bmatrix} X \\ &+ \begin{bmatrix} 0.4309 & 0.0000 & -3.6673 \\ -3.8202 & 0.0000 & -9.7782 \\ -6.3684 & 0.0000 & -4.5933 \end{bmatrix} = 0. \end{aligned} \tag{8}$$

By using Newton's method, we get the solvent

$$X = \begin{bmatrix} -1.06920 & 0.3590 & -0.9020 \\ -0.2692 & 0.1887 & -0.4248 \\ 1.5866 & -2.0719 & 0.9482 \end{bmatrix}$$

of $Q_2(X)$. Starting Newton's method with $X_0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$, we have the results displayed in Figure 2. In this case, we expect that the condition number of this problem is bad. That is, it is meaningful to calculating the condition number.

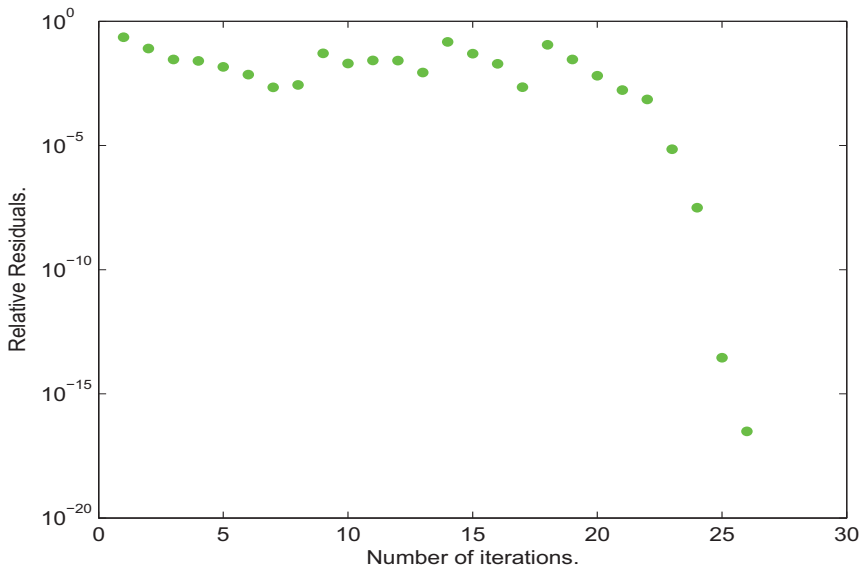

Figure 2. Convergence for the problem (8)

The following example, we consider the results of normwise condition number.

Example 4.3. *The quadratic matrix equation $Q_3(X)$ with*

$$A_3 = \begin{bmatrix} 10^4 & -1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, B_3 = \begin{bmatrix} 0.5 & 1 & 0 \\ 0 & 0.5 & 1 \\ 10^4 & 0 & 0.5 \end{bmatrix},$$

and

$$C_3 = 10^4 \begin{bmatrix} -3.3333 & -0.0002 & -0.0003 \\ -0.0007 & -0.0008 & 0.0003 \\ -0.0003 & 0 & -5.0003 \end{bmatrix}.$$

In this example, $\alpha = \|A_3\|_F, \beta = \|B_3\|_F, \gamma = \|C_3\|_F$.

$\kappa_1(\varphi)$	$1.4427e+04$
$\kappa_2^U(\varphi)$	$2.4988e+04$
$\kappa_2^M(\varphi)$	$2.5948e+04$
$\kappa_3(\varphi)$	$4.5497e+04$

Table 3. Condition numbers of $Q_3(X)$

From Table 3, we can see that $\kappa_3(\varphi)$ is the largest normwise condition number.

5. Conclusion

In this section, we give a summary of our works and compare numerical experimental results. The quadratic matrix equation in (1) arises in some applications. We introduced two conditioning analysis of quadratic matrix equation. And we presented three kinds of normwise condition numbers.

References

[1] G. J. Davis, *Numerical solution of a quadratic matrix equation*, SIAM J. Sci. Stat. Comput., **2** (1981), 164–175.
[2] I. Gohberg and I. Koltracht, *Mixed, componentwise, and structured condition numbers*, SIAM J. Matrix Anal. Appl., **14** (1993), 688–704.
[3] N. J. Higham and Hyum-Min Kim, *Solving a quadratic matrix equation by Newton's method with line searches*, SIAM J. Matrix Anal. Appl., **23** (2001),, 303–316.
[4] R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, (1995).
[5] Yiqin Lin and Yimin Wei, *Condition Numbers of the Generalized Sylvester Equation*, IEEE Transactions on Automatic Control, **52** (2007), 2380–2385.
[6] R. D. Skeel, *Scaling for numerical stablity in Gaussian elimination*, J. Assoc. Comput. Mach., **26** (1979), 494–526.

Hye-Yeon Kim
Department of Mathematics, Pusan National University, Busan, 609-735, Republic of Korea
E-mail address: hyeyeon@pusan.ac.kr

Hyun-Min Kim
Department of Mathematics, Pusan National University, Busan, 609-735, Republic of Korea
E-mail address: hyunmin@pnu.edu