Matrix Quadratic Equations: Algebraic Geometry

Rama Garimella, Alexander Rumyantsev and B K Krishna Kumar

March 4, 2020

Matrix Quadratic Equations: Algebraic Geometry

Garimella Rama Murthy*, Alexander Rumyantsev**, B. Krishna Kumar***

*Department of Computer Science, Mahindra Ecole Centrale, Hyderabad, India

**Karelian Research Center, Russian Academy of Sciences, Russia

***Department of Mathematics, Anna University, Chennai, India

ABSTRACT

In this research paper, we consider matrix quadratic equations in which the coefficient matrices as well as unknown matrix are 2 x 2 matrices. It is shown that linear algebraic techniques enable complete characterization of matrix solutions of such equations. Specifically such matrix equations arising in queueing theory are explicitly studied.

1. INTRODUCTION:

The development of algebraic symbolism was, in part, motivated by the concepts: zero, negative numbers. Solutions of linear algebraic equations in one variable with coefficients being rational numbers lead to the concept of rational numbers. As a natural generalization, solving quadratic equations was attempted by mathematicians across the planet. Indian mathematicians solved quadratic equations using the "completion of square" technique. Efforts to solve higher degree polynomial equations led to the research area of "group theory".

Mathematicians such as Gauss attempted solving a system of linear equations in multiple variables leading to the research area of "linear algebra". Using the method of elimination, Gauss successfully solved system of linear equations (so called "Gaussian Elimination"). As a natural generalization, polynomial equations with matrix coefficients and single matrix unknown are attempted for solution. Mathematicians proved interesting theorems related to the solution of matrix polynomial equations.

Bezout proved an interesting theorem related to multi-variate polynomial equations of finite degree. This theorem was a central contribution to the research area of "algebraic geometry". It was realized by the authors that matrix polynomial equation in a single matrix unknown represents a structured system of multi-variate polynomial equations. Thus, determination of their solutions is a contribution to algebraic geometry. In [RaA], the authors showed that solution of a matrix quadratic equation, with unknown matrix as well as coefficient matrices being 2 x 2 matrices, can be determined by a formula involving coefficient matrices under some conditions. Spcifically, such a structured matrix quadratic equation arising in queueing theory was considered and one of its matrix zeroes was determined by a formula involving coefficient matrices and its eigenvalues. Such a result motivated us to study arbitrary matrix quadratic equations in which the unknown matrix as well as coefficient matrices are 2 x 2 matrices. The results of such an effort are documented in this research paper.

2. MATRIX QUADRATIC EQUATIONS: 2 X 2 COEFFICIENT MATRICES: SOLUTIONS:

Consider an arbitrary matrix quadratic equation in which the unknown matrix, X as well as the coefficient matrices $\{ B_0, B_1, B_2 \}$ are 2 x 2 matrices i.e.

$$X^2 B_2 + X B_1 + B_0 \equiv \bar{0} \quad \ldots\ldots\ldots\ldots\ldots\ldots. (1).$$

Let the unknown matrix, X *and* $B_2,\ B_1,\ B_0$ be given by

$$X = \begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix}, \quad B_2 = \begin{bmatrix} b_{11}^{(2)} & b_{12}^{(2)} \\ b_{21}^{(2)} & b_{22}^{(2)} \end{bmatrix}, \quad B_1 = \begin{bmatrix} b_{11}^{(1)} & b_{12}^{(1)} \\ b_{21}^{(1)} & b_{22}^{(1)} \end{bmatrix}, B_0 = \begin{bmatrix} b_{11}^{(0)} & b_{12}^{(0)} \\ b_{21}^{(0)} & b_{22}^{(0)} \end{bmatrix}.$$

The above matrix quadratic equation corresponds to the following system of 4-variate equations.

$$(x_1^2 + x_2 x_3)\, b_{11}^{(2)} + (x_1 x_2 + x_2 x_4)\, b_{21}^{(2)} + (x_1)\, b_{11}^{(1)} + (x_2)\, b_{21}^{(1)} + b_{11}^{(0)} = 0.$$

$$(x_1^2 + x_2 x_3)\, b_{12}^{(2)} + (x_1 x_2 + x_2 x_4)\, b_{22}^{(2)} + (x_1)\, b_{12}^{(1)} + (x_2)\, b_{22}^{(1)} + b_{12}^{(0)} = 0.$$

$$(x_3 x_1 + x_4 x_3)\, b_{11}^{(2)} + (x_3 x_2 + x_4^2)\, b_{21}^{(2)} + (x_3)\, b_{11}^{(1)} + (x_4)\, b_{21}^{(1)} + b_{21}^{(0)} = 0.$$

$$(x_3 x_1 + x_4 x_3)\, b_{12}^{(2)} + (x_3 x_2 + x_4^2)\, b_{22}^{(2)} + (x_3)\, b_{12}^{(1)} + (x_4)\, b_{22}^{(1)} + b_{22}^{(0)} = 0.$$

Thus, the problem of finding unknown matrix falls in the research area of algebraic geometry. But mathematicians capitalized the fact that the system of equations are structured and proved interesting results. Specifically, linear algebraic tools are utilized to determine the unknown matrix solutions of equation (1). We now summarize the well known results [Gan].

The following lemma enables determination of all possible eigenvalues of unknown matrix solutions of equation (1). We denote the following result as "Factorization Lemma". Such lemma applies to matrix quadratic equations, where the square unknown matrix as well as coefficient matrices are of arbitrary dimension.

Lemma 1: $(\mu^2 B_2 + \mu B_1 + B_0) \equiv (\mu I - X)(\mu B_2 + X B_2 + B_1)$. Thus, we have that

$$f(\mu) = Det\,(\mu^2 B_2 + \mu B_1 + B_0) = Det\,(\mu I - X)\, Det\,(\mu B_2 + X B_2 + B_1).$$

Thus, all possible eigenvalues of unknown matrix solutions are a subset of zeroes of the determinental polynomial $f(\mu) = Det\,(\,\mu^2 B_2 + \mu B_1 + B_0)$.

Proof: The result follows by showing that RHS and LHS agree. Details are avoided for brevity.

Note: $f(\mu)$ is a polynomial of degree 2 N (where $\{\,X, B_i^{\prime}s\,\}$ are N x N square matrices). If the coeficients matrices i.e. $B_i^{\prime}s$ are upper/lower triangular matrices, the eigenvalues of all solutions can be determined by solving scalar quadratic equations. Once the eigenvalues are known, the following theorem enables determination of unknown matrix X (with those eigenvalues) as a solution of homogeneous linear systems of equations.

Theorem 1: Consider an arbitrary matrix quadratic equation of the form in equation (1). Let the dimension of X be N. Then all possible solutions of (1) are divided into atmost $\binom{2N}{N}$ equivalence classes (equivalence classes are specified based on same set of Eigenvalues) and solution in each class is determined as the solution of a linear system of equations.

Proof: Refer [Gan].

- The above results are applicable to unknown matrix (as well as coefficient matrices) of arbitrary dimension 'N'.

- Now we consider the case where N = 2. In this case $f(\mu)$ is a polynomial in $'\mu'$ of degree '4'. Hence, its zeroes can be explicitly determined by algebraic formulae in its coefficients. Thus, all possible eigenvalues of unknown matrix solutions can be determined by algebraic formulae.

- Now, we reason that in this case (N = 2), the unknown matrix can be expresses by an algebraic formula involving coefficient matrices and its eigenvalues. Details are provided below.

  Let $\{\,\alpha, \beta\,\}$ be the eigenvalues of 2 x 2 unknown matrix X, and let its characteristic polynomial $g(\mu\,)$ be given by

  $$g(\mu\,) = \;(\,\mu - \alpha\,)\,(\,\mu - \beta\,) = \;\mu^2 - \mu\,(\,\alpha + \beta\,) + \;\alpha\beta = \;\mu^2 + \;b_1\;\mu \;+\; b_0\;, where$$
  $$b_0 = Det(X)\;\; and\;\; b_1 = -Trace(X).$$

  By Cayley-Hamilton theorem, we have that

  $$g(X) = \;X^2 + \;b_1\;X \;+\; b_0\,I \equiv \bar{0}\,.$$

  Hence,

$$X^2 = -b_1 X - b_0 I.$$

Thus, substituting in the matrix quadratic equation and simplifying, we have that

$$X(B_1 - b_1 B_2) + (B_0 - b_0 B_2) \equiv \bar{0}.$$

$Hence, if \ (B_1 - b_1 B_2) \ is \ non-singular, then$

$$X = -(B_0 - b_0 B_2)(B_1 - b_1 B_2)^{-1}.$$

$But, if \ (B_1 - b_1 B_2) \ is \ singular, \ then$ there are infinitely many solutions of the matrix quadratic equation (1) with $\{\alpha, \beta\}$ as the eigenvalues.

Now, we identify conditions under which $(B_1 - b_1 B_2)$ is singular. It can be readily seen that

$Det((B_1 - b_1 B_2) = b_1^2 \, Det(B_2) + b_1 \theta + Det(B_1),$ $\quad where \ '\theta' \ is \ expressed$

In terms of elements of $\{B_1, \ B_2\}$. *Thus, there are atmose '2' values of Trace(X) for which* $(B_1 - b_1 B_2)$ *is singular. These values of Trace(X) can be real values or complex numbers.*

We know that there are exactly $\binom{4}{2} = 6$ *pairs of zeroes of* $g(\mu)$.

Hence, we readily infer that the number of UNIQUE solutions of (1) can be determined by the following equation:

Number of Unique Solutions of (1) $\geq Maximum\{Number \ of \ Distinct \ Trace \ Values - 2, 0\}$.

***Note**: There are 6 possible trace values and some of them could be equal.*

***Note:** Given a pair of zeroes of* $g(\mu)$ *that are potential eigenvalues of a solution X (of (1)), either a unique X, exists or infinitely many solutions exist (with those pair of eigenvalues).*

$\boldsymbol{Note}$: *Suppose* $\{B_0, B_1, B_2\}$ *are matrices with real valued components. Then, it readily follows that the zeroes of* $g(\mu)$ *occur in complex conjugate pairs.*

- From, the above discussion, when $B_i's$ are all 2 x 2 matrices, using the Cayley-Hamilton Theorem, we infer that $X_j's \ for \ j \geq 2$ (with a fixed pair of eigenvalues of X) can be expressed in terms of matrices $\{X, I\}$ using scalar coefficients determined by the coefficients of characteristic polynomial of X. For instance, we readily have that

  $$X^3 = (b_1^2 - b_2)X + b_1 b_2 I.$$

  $Now \ letting, \ b_1^{(2)} = b_1, \ b_2^{(2)} = b_2 \ and \ b_1^{(3)} = (b_1^2 - b_2), b_2^{(3)} = b_1 b_2$, we have that the following recursive equation holds:

  $$\begin{bmatrix} -b_1^{(3)} \\ -b_2^{(3)} \end{bmatrix} = \begin{bmatrix} -b_1^{(2)} & 1 \\ -b_2^{(2)} & 0 \end{bmatrix} \begin{bmatrix} -b_1^{(2)} \\ -b_2^{(2)} \end{bmatrix} = C_{(2)} \begin{bmatrix} -b_1^{(2)} \\ -b_2^{(2)} \end{bmatrix}, \quad where \ C_{(2)} \ is \ a \ companion \ matrix.$$

  This recursion was first observed in [RaA2]. By letting,

$$X^4 = -b_1^{(4)} X - b_2^{(4)} I$$

It can be readily shown that

$$\begin{bmatrix} -b_1^{(4)} \\ -b_2^{(4)} \end{bmatrix} = \left(C_{(2)}\right)^2 \begin{bmatrix} -b_1^{(2)} \\ -b_2^{(2)} \end{bmatrix}.$$

In general, X^M *can be expressed, in terms of* $\{X, I\}$ *with two suitable coefficients*. The two coefficients are obtained using the following recursion (generalization of above result):

$$\begin{bmatrix} -b_1^{(M)} \\ -b_2^{(M)} \end{bmatrix} = \left(C_{(2)}\right)^{M-2} \begin{bmatrix} -b_1^{(2)} \\ -b_2^{(2)} \end{bmatrix}.$$

Thus, coefficients of higher powers of X can be expressed in terms of a 2 x 2 companion matrix and the coefficients of characteristic polynomial of X.

Note: This result can easily be generalized to higher degree matrix polynomial equations and even matrix power series equations using the Cayley-Hamilton Theorem.

- We now now consider a solution X of (1) and arrive at another related matrix, H which satisfies a dual matrix quadratic equation of the following form:

$$B_2 H^2 + B_1 H + B_0 \equiv \bar{0} \quad \ldots\ldots\ldots\ldots\ldots\ldots (2).$$

Suppose B_2 *is non* $-$ *singular.*

Let us define the matrix H as $H = -B_2^{-1}(X B_2 + B_1)$. It readily follows that

$$H^2 = B_2^{-1}(X B_2 + B_1)\, B_2^{-1}(X B_2 + B_1)$$

$$H^2 = B_2^{-1}(X + B_1 B_2^{-1})(X B_2 + B_1)$$

$$B_2 H^2 = (X + B_1 B_2^{-1})(X B_2 + B_1) = X^2 B_2 + X B_1 + B_1 B_2^{-1}(X B_2 + B_1)$$

Using the fact that X *is a solution of* (1), *we have*

$$B_2 H^2 = -B_0 - B_1 H.$$

Hence, it readily follows that H *satisfies the following matrix equation*

$$B_2 H^2 + B_1 H + B_0 \equiv \bar{0}\,.$$

Thus, the solutions X, H *satisfy dual matrix quadratic equations*. It readily follows that given solution H, we can obtain X in the following manner:

$$X = -(B_2 H + B_1)\, B_2^{-1}.$$

Note: *In queueing theory such a matrix ,* H *naturally arises.*
We briefly discuss this issue in section 3.

3. MATRIX QUADRATIC EQUATIONS: QUEUEING THEORY:

Structured matrix polynomial equations naturally arise in the equilibrium analysis of a class of Markov chains (in discrete time as well as continuous time) called G/M/1-type Markov processes as well as M/G/1-type Markov processes. Among them, in the equilibrium analysis of Quasi-Birth-and-Death (QBD) processes, structured matrix quadratic equations of the following form naturally arise:

$$R^2 A_2 + R A_1 + A_0 \equiv \bar{0} \ldots\ldots (3), \quad where$$

R *is called the 'rate matrix' and constitutes the minimal non − negative solution of the above structured matrix quadratic equation ('minimal' in the sense that* the sum of *all elements of the matrix is minimum*). $\{A_0, A_2\}$ are non-negative matrices and the matrix A_1 is diagonally dominant with negative diagonal elements and non-negative off-diagonal elements. Further

$A = A_0 + A_1 + A_2$ *is a generator matrix i.e. digonal elements of 'A' are negative and off − diagonal elements are non − negative and all the rowsums (i.e. sum of all rowwise elements) are zero.*

Thus, it readily follows that factorization lemma and Theorem 1 readily apply to the solutions of above structured matrix quadratic equation (i.e. equation (3) . In fact, all the results discussed in Section 2, naturally apply. In [Rama1, RaKC], computation of Jordan Canonical form of rate matrix, R is discussed in complete detail. We now study issues related to all other matrix solutions of (3). In that effort, the following lemma related to the zeroes of $g(\mu) = Det\,(\mu^2 A_2 + \mu A_1 + A_0)$ readily **follows.**

Lemma 2:

All the zeroes of $g(\mu)$ (i.e. eigenvalues of all possible solutions of equation (3)) are distinct. Hence all matrix solutions of (3) are diagonalizable.

Proof: It is well known [Neu] that the spectral radius of rate matrix 'R' is strictly less than one. Since 'R' is an irreducible non-negative matrix, by Perron-Frobenius Theorem, the spectral radius is real, positive, simple and the corresponding left/ right eigenvector has strictly positive components.

Since, 'R' has real valued components, the trace (R) is a real number. Hence, the other eigenvalue of R is real and strictly less than spectral radius τ . Let us label, the smaller eigenvalue or 'R' as 'α'.

By factorization lemma, we have that

$$(\mu^2 A_2 + \mu A_1 + A_0) \equiv (\mu I - R)(\mu A_2 + R A_2 + A_1).$$

Hence

$(A_2 + A_1 + A_0)\,\bar{e} = \bar{0}$ (where $\bar{e}$ is a column vector all of whose components are '1'), since $A = (A_2 + A_1 + A_0)$ is a generator matrix. Hence '1' is a zero of $g(\mu)$, a 4^{th} degree polynomial. The following lemma deals with the remaining zero Q.E.D.

Now, we reason in the following lemma that the other remaining zero of $g(\mu)$ is strictly larger than one. Let such zero be denoted by $'\delta'$.

Lemma 3: There are two distinct zeroes of $g(\mu)$ that are on or outside unit circle

Proof: In the above lemma, we reasoned that '1' is a zero of $g(\mu)$. Using factorization lemma with $\mu = 1$, we have that

$$A = (A_2 + A_1 + A_0) \equiv (I - R)(A_2 + R\,A_2 + A_1).$$

Since 'A' is a generator matrix, we have that $(A_2 + A_1 + A_0)\,\bar{e} = \bar{0}$. Using the fact that the spectral radius of irreducible rate matrix, R is strictly less than one, we have that

$$(A_2 + R\,A_2 + A_1)\,\bar{e} = \bar{0} = R\,A_2\,\bar{e} + (A_2 + A_1)\,\bar{e}.$$

Hence, it follows that

$$R\,A_2\,\bar{e} = A_0\bar{e}\ .$$

Since, A_1 is diagonally dominant and $(A_1)\,\bar{e} = -(A_0 + A_2)\,\bar{e}$, $(R\,A_2 + A_1)$ is strictly diagonally dominant since $A_2\,\bar{e} > \bar{0}$. Hence $(R\,A_2 + A_1)$ is non-singular. Further

$$(R\,A_2 + A_1)^{-1}\,(A_2 + R\,A_2 + A_1)\,\bar{e} = \bar{0}\,.$$

Thus, we have that

$$-(R\,A_2 + A_1)^{-1}\,A_2\,\bar{e} = \bar{e}.$$

Also, since $(R\,A_2 + A_1)$ is strictly diagonally dominant (with negative diagonal elements and non-negative off-diagonal elements), $-(R\,A_2 + A_1)^{-1}$ is a non-negative matrix. Hence $-(R\,A_2 + A_1)^{-1}\,A_2$ is a stochastic matrix with spectral radius ONE. Hence, if μ is an eigenvalue with the corresponding right eigenvector $\bar{f}$, we have that

$$-(R\,A_2 + A_1)^{-1}\,A_2\,\bar{f} = \mu\,\bar{f}\ \ with\ \ |\mu| < 1.$$

Thus, $(\mu(R\,A_2 + A_1) + A_2)\,\bar{f} = \bar{0}$ for every eigenvalue μ. Let $\frac{1}{\mu} = \theta$.

Hence, it readily follows that

$$((R\,A_2 + A_1) + \theta\,A_2)\,\bar{f} = \bar{0}.$$

Hence all the zeroes of $g(\mu)$, other than those of rate matrix R are all on or outside the unit circle. There is exactly one zero lying at '1'. Q.E.D.

Note: The above proof is more general and applies to the case where the dimension of coefficient matrices (i.e A_0, A_1, A_2) is an arbitrary integer value N (not just N=2).

Note: From the above proof, it is clear that the eigenvalues of the stochastic matrix $-(R A_2 + A_1)^{-1} A_2$ are '1', '$\mu = \frac{1}{\delta}$, where δ is a real number strictly larger than 1.

Uniqueness of Solutions of Rate Matrix based Matrix Quadratic Equation:

- From the above discussion, it is clear that all the four zeroes of $g(\mu)$ *are real and distinct*. Specifically $\alpha < \tau < 1 < \delta$.

- Hence, all 6 possible trace values are

$$\{ \alpha + \tau, \alpha + 1, \alpha + \delta, \tau + 1, \tau + \delta, 1 + \delta \}.$$

  In view of the above discussion on the values of four zeroes of $g(\mu)$, the following inequalities hold true

$$\alpha + \tau < \alpha + 1 < \alpha + \delta < \tau + \delta < 1 + \delta.$$

  Thus, there are 5 distinct values of trace of potential matrix solutions. Hence based on earlier reasoning (equation ( )), there are atleast 3 UNIQUE matrix solutions of the structured matrix quadratic equation arising in the equilibrium analysis of Quasi-Birth-and-Death process.

Using an alternative reasoning, we prove that for the trace value of $(\alpha + 1)$, the associated matrix solution is UNIQUE i.e. we essentially reason that $(A_1 + (1 + \alpha) A_2)$ is non-singular

Lemma 4: $(A_1 + (1 + \alpha) A_2)$ is strictly diagonally dominant and hence is non-singular

Proof: $A_1 + (1 + \alpha) A_2 = A_1 + A_2 + \alpha A_2$. Also, we readily know that

$(A_2 + A_1 + A_0) \bar{e} = \bar{0}$ (where $\bar{e}$ is a column vector all of whose components are '1').

Hence, it is sufficient to show that $\alpha A_2 \bar{e} < A_0 \bar{e}$, *where the inequality holds for all the components of the vectors*. From Lemma (), it is clear that

$R A_2 \bar{e} = A_0 \bar{e}$ *with* $R = \alpha E_1 + \tau E_2$, *where* E_1, E_2 *are residue matrices such that* $E_1 + E_2 = I$ *i.e. identity matrix (all components of* E_2 *are positive by Perron's Theorem).*

Thus, $R > \alpha E_1 + \alpha E_2$ (componentwise inequality). Equivalently $R > \alpha I$.

Since A_2 *is a non-negative matrix, we have that* $\alpha A_2 \bar{e} > A_0 \bar{e}$.

Since A_1 is a diagonally dominant matrix with negative diagonal elements and non-negative off-digonal elements, from the above discussion, it readily follows that $(A_1 + (1 + \alpha) A_2)$ is strictly diagonally dominant and hence is non-singular Q.E.D.

Corollary: Suppose $A_2 = A_0$. Using the same reasoning, it follows that $(A_1 + (1+\tau)A_2)$ is strictly diagonally dominant and hence is non-singular.

Lemma 5: Let the zeroes of determinental polynomial $g(\mu) = Det\,(\mu^2 A_2 + \mu A_1 + A_0)$ be $\{ \alpha, \tau, 1, \delta \}\; with\; \alpha < \tau < 1 < \delta \;\; and \;\; \tau > 0$. Then

$$(\alpha)(\tau)(\delta) \; = \; \frac{Det(A_0)}{Det\,(A_2)} \, .$$

Thus, the sign/polarity of $\{ \alpha, \delta \}$ are suitably constrained.

Proof: From Lemma 3, we have that $-(R\, A_2 + A_1)^{-1}\, A_2 = P$ is a stochastic matrix with eigenvalues being $\{ 1, \frac{1}{\delta} \}$. Hence, we have that (Considering 2 x 2 matrix case)

$$Det(P) = \frac{Det(A_2)}{Det(RA_2 + A_1)} = \; (1) \left(\frac{1}{\delta}\right).$$

Since, the rate matrix satisfies the matrix quadratic equation, we have that

$$R\;(R\; A_2 + A_1) = \; -A_0.$$

Since, we consider the case of 2 x 2 matrices, we have that

$$Det(R\; A_2 + A_1) \;\; = \;\; \frac{Det(A_0)}{Det(R)}\, .$$

But, $Det(R) = (\alpha)(\tau)$. Using the above equations, we have that

$$\delta = \frac{Det\;(\;R\; A_2 + A_1)}{Det(\; A_2)} = \frac{Det(A_0)}{Det(R)} \;.\frac{1}{Det(A_2)} = \;\frac{Det(A_0)}{Det(A_2)} \,.\frac{1}{(\alpha)(\tau)} \;.$$

Hence, we readily infer that

$$(\alpha)(\tau)(\delta) \;\; = \;\; \frac{Det(A_0)}{Det(A_2)}\, .$$

From the above equation, the following inferences readily follow:

Case I: If both $Det(A_0)$, $Det(A_2)\; are\;\; of\;\; the\;\; same\;\; sign$, then $\{ \alpha, \delta \}$ are both of the same sign. Hence, if $\{ \alpha, \delta \}\; are\;\; both\;\; positive$, then all the solutions of 2 x 2 matrix based quadratic equation are all positive definite.

Case II: : If both $Det(A_0)$, $Det(A_2)\; are\;\; of\;\; the\;\; opposite\;\; sign$, then $\{ \alpha, \delta \}$ are of the opposite sign. Q.E.D.

FUTURE RESEARCH WORK:

In future versions of this preprint, we propose to document our results related to the following topics.

4. RELATED MATRIX QUADRATIC EQUATIONS: SOLUTIONS:

We now consider the "dual" of matrix quadratic equation discussed in Section 2. i.e. $B_2Y^2 + B_1Y + B_0 \equiv \bar{0}$.

In the case of such matrix quadratic equation, the "$factorization\ lemma$" takes the following form:

Lemma 6: The eigenvalues of all solutions of ( ) are zeroes of the determinental polynomial

$$f(\mu) = Det\,(\mu^2 B_2 + \mu B_1 + B_0)\,.$$

$\boldsymbol{Proof}$: $The\ following\ identity\ readily\ follows$:

$$\mu^2 B_2 + \mu B_1 + B_0 \equiv (\mu\, B_2 + Y B_2 + B_1)(\mu\, I - Y).$$

$Taking\ the\ determinent\ on\ both\ the\ sides, we\ have\ that$

$$Det(\mu^2 B_2 + \mu B_1 + B_0) = Det(\mu\, B_2 + Y B_2 + B_1)\, Det(\mu\, I - Y).$$

$Hence\ the\ claim\ readily\ follows. \qquad Q.E.D.$

5. ALGEBRAIC RICCATI EQUATIONS: MATRIX QUADRATIC EQUATIONS:

Quadratic matrix equations of the form

$$XCX - AX - XD + B \equiv \bar{0}$$

are called Nonsymmetric Algebraic Riccati Equations (NARE) (in which X is an *m x n* unknown matrix and the coefficient matrices *A, B, C, D* have dimensions *m x m, m x n, n x m and n x n respectively).* For the purposes of simplicity, we assume that *m = n [BBMP].*

Also, continuous time algebraic Riccati equation is of the form:

$$XCX - AX - XA^T + B \equiv \bar{0}\,,\ where\ B\ and\ C\ are\ symmetric\ matrices.$$

We now consider NARE with (m= n) and *A = -D*. Thus, we have that

$$XCX - (AX + XD) + B = XCX + B \equiv \bar{0}.$$

$Multiplying\ on\ both\ sides\ of\ the\ above\ equation\ by\ matrix\ 'C', \quad we\ have\ that$

$$XCXC + BC = (XC)^2 + BC \equiv \bar{0}.$$

$$Thus, \quad XC = Matrix\ Sqare\ Root\ of\ (-B\,C).$$

$If\ B\,C\ is\ negative\ definite\ matrix\ and\ C\ is\ non-singular, \quad we\ have\ that$

$$X = [\,Matrix\ Square\ Root\ of\ (-B\,C)\,]\ C^{-1}.$$

In general, there will be multiple solutions for the matrix square root of the matrix (-B C) if it is not a positive definite matrix.

$\boldsymbol{Note}$: $The\ above\ idea\ of\ utilizing\ matrix\ square\ root\ arose\ in\ a\ discussion\ with\ Prof. Dario\ Bini.\ The\ following\ discussion\ is\ a\ generalization.$

Suppose in NARE, $A = D = I$. Then we have that

$$X\,C\,X - 2\,X + B \equiv \bar{0}.$$

Letting $X\,C = Y$, *we have that*

$$Y^2 - 2\,Y + B\,C \equiv \bar{0}.$$

Such a matrix quadratic equation can be readily solved using the well known techniques developed for arbitrary matrix quadratic equations [*Gan*].

6. GENERALIZATION TO MATRIX POLYNOMIAL EQUATIONS:

We reason that most of the results in earlier sections readily generalize for more general matrix polynomial equations. In the following, one such result in Section 2 is generalized.

In Section 2, using Cayley-Hamiton Theorem, it was shown (when the *unknown matrix* X *and coefficient matrices i.e.* $B_i's$ *are* $2\,x\,2$ *matrices*) that higher powers of X *i.e.* X^m *for* $m \geq 3$ can be expressed in terms of linear combination of $\{I, X\}$. The coefficients in the linear combination are determined using a 2 x 2 companion matrix (whose elements are the coefficients of characteristic polynomial of X) and the coefficients of characteristic polynomial of X. This result naturally generalizes for higher degree matrix polynomial equation.

Let us consider an M^{th} *degree matrix polynomial equation in which the unknown matrix, as well as coefficient matrices are* $N\,x\,N$ *matrices.* Let the coefficients of *characeristic polynomial of* X *be denoted by* $b_{(N)}^{(1)}$ *and* $C_{(N)}$ *be the companion matrix whose elements are determined by the characteristic polynomial of* X.

By Cayley–Hamilton Theorem, it is clear that higher powers of X *i.e* $\{X^j : j \geq N\}$ Can be expressed in terms of linear combination of matrices $\{I, X, X^2, \ldots, X^{N-1}\}$. Thus, let the coefficients in the linear combination be the elements of a vector denoted by

$b_{(N)}^{(m)}$: *i.e. coefficients in linear combination leading to* $X^{(m)}$ (*in terms of* $\{I, X, \ldots, X^{N-1}\}$)

As in Section 2, it can easily be shown that

$$b_{(N)}^{m} = C_{(N)}^{m-N}\, b_{(N)}^{(1)} \;\; for\;\; m \geq N \;\; with\; C_{(N)} \;\; being\;\; N\,x\,N \;\; companion\;\; matrix.$$

7. CONCLUSIONS:

In this research paper, we provide interesting results related to solving matrix quadratic equations in which the coefficient matrices as well as unknown matrix are 2 x 2 matrices. We readily realize that the results can easily be generalized to such matrix polynomial equations of arbitrary degree.

REFERENCES:

[BBMP] Dario Bini, Bruno, Meini, Poloni "Nonsymmetric algebraic Riccati equations associated with an M-matrix: recent advances and algorithms," Proceedings of Numerical Methods for Structured Markov Chains, 2007

[Gan] F. Gantmacher, "Theory of Matrices," Chelsea Publisher, AMS.

[Neu] M. F. Neuts,"Matrix Geometric Solutions in Stochastic Models," Johns Hopkins University Press, Baltimore, 1981

[Rama1] Rama Murthy, G," Transient and equilibrium analysis of computer networks: Finite memory and matrix geometric recursions," PhD. Thesis, Purdue University, West Lafayette (1989)

[RaKC] Rama Murthy, G., Kim, M., Coyle, E. J." Equilibrium analysis of skip-free Markov chains: Non-linear Matrix Equations," Communications in Statistics – Stochastic Models. 4, pp. 547–571 (1991)

[**RaA1**] G. Rama Murthy and Alexander Rumyantsev**,"** On an exact solution of the rate matrix of Quasi-Birth-Death process with small number of phases" **Proceedings of 31st European Conference on Modelling and Simulation (ECMS 2017), Budapest, Hungary .**

[RaA2] G. Rama Murthy and Alexander Rumyantsev," On an exact solution of the rate matrix of G/M/1-type Markov Process with small number of phases**," Journal of Parallel and Distributed Computing, Elsevier Science, Vol.119, September 2018, pages 172-178………JOURNAL PAPER**