POLYNOMIAL EQUATIONS OVER MATRICES

Robert Lee Wilson

Here are two well-known facts about polynomial equations over the complex numbers ($\mathbf{C}$):

(I) (Vieta Theorem) For any complex numbers $x_1, ..., x_n$ (not necessarily distinct), there is a unique monic polynomial over $\mathbf{C}$

$$f(x) = x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0$$

$$= (x - x_1)...(x - x_n)$$

such that the equation $f(x) = 0$ has roots $x_1, ..., x_n$ (counting multiplicities). Then

$$a_{n-i} = (-1)^i \sum_{j_1 < ... < j_i} x_{j_1} ... x_{j_i}.$$

(II) (Fundamental Theorem of Algebra) The equation over $\mathbf{C}$

$$x^n + a_{n-1}x^{n-1} + ... + a_1 x + a_0 = 0$$

has a root. Consequently, it has exactly n roots (counting multiplicities).

This article describes the similar statements that can be made about a polynomial equation

$$X^n + A_{n-1}X^{n-1} + ... + A_1 X + A_0 = 0$$

over $M_k(\mathbf{C})$, the algebra of k by k matrices over $\mathbf{C}$. For simplicity I will assume $n = 2$ and $k = 2$. All the results can be extended to general n and k, but the notation gets more complicated. Such problems for polynomials of low degree (particularly for $n = 2$) have been treated by several authors since they arise naturally in control theory and in functional analysis. See, for example, [GHR,LR]. For general n, the solution of the diagonalizable case is due to Fuchs and Schwarz [FS].

The differences between the situation for equations over the complex numbers and the situation for equations over matrices arise for two reasons: multiplication in $M_k(\mathbf{C})$ is not commutative and not all matrices in $M_k(\mathbf{C})$ are invertible.

Analogues of the Vieta Theorem

Let X_1 and X_2 be roots of the quadratic equation $X^2 + A_1X + A_0 = 0$ over any algebra (not necessarily commutative - e.g., $M_k(\mathbf{C})$).

Then we have

$$X_1^2 + A_1X_1 + A_0 = 0$$

and

$$X_2^2 + A_1X_2 + A_0 = 0.$$

Taking the difference gives

$$X_1^2 - X_2^2 + A_1(X_1 - X_2) = 0.$$

Replace

$$X_1^2 - X_2^2$$

by

$$X_1(X_1 - X_2) + (X_1 - X_2)X_2.$$

(This is the noncommutative version of the well-known formula $u^2-v^2 = (u+v)(u-v)$ that holds in the commutative case.) We obtain

$$X_1(X_1 - X_2) + (X_1 - X_2)X_2 + A_1(X_1 - X_2) = 0.$$

Thus

$$-A_1(X_1 - X_2) = X_1(X_1 - X_2) + (X_1 - X_2)X_2. \tag{1}$$

This has two consequences:

First, if $X_1 - X_2$ is invertible we may multiply on the right by its inverse and obtain

$$-A_1 = X_1 + (X_1 - X_2)X_2(X_1 - X_2)^{-1}.$$

To make this look nicer, write

$$y_1 = X_1$$

and

$$y_2 = (X_1 - X_2)X_2(X_1 - X_2)^{-1}.$$

Then

$$-A_1 = y_1 + y_2$$

and

$$A_0 = -X_1^2 - A_1X_1 = -y_1^2 + y_1^2 + y_2y_1.$$

In other words:

$$A_1 = -(y_1 + y_2)$$

and

$$A_0 = y_2y_1.$$

These are close analogues of the formulas in the commutative case, but they involve rational functions, not just polynomials. This generalization of the Vieta Theorem (which can be extended to equations of degree n by using the theory of quasideterminants) is due to Gelfand and Retakh [GR1,GR2].

Second, equation (1) may be rewritten as

$$-(X_1 + A_1)(X_1 - X_2) = (X_1 - X_2)X_2. \tag{2}$$

Now work in $M_2(\mathbf{C})$ and take

$$X_1 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, X_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Then

$$(X_1 - X_2)X_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

But the rows of the matrix

$$-(X_1 + A_1)(X_1 - X_2)$$

must be contained in the row space of the matrix

$$X_1 - X_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Thus it is impossible to satisfy (2) and hence there is no quadratic equation over $M_2(\mathbf{C})$ with roots

$$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Solutions of polynomial equations- how many solutions can there be?
Some examples:

(1) $X^2 - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = 0$ has no solution.

To see this, suppose that $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is a solution and observe that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^2 = \begin{bmatrix} a^2 + bc & b(a+d) \\ c(a+d) & bc + d^2 \end{bmatrix}.$$

If this is equal to $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, then $a + d \neq 0$. But $c(a+d) = 0$, so $c = 0$. But then (comparing the entries on the diagonal) $a^2 = d^2 = 0$, so $a = d = 0$, a contradiction.

We will return to this example later and see another, less computational, way to analyze it.

(2) $X^2 = 0$ has infinitely many solutions

For example, $\begin{bmatrix} 0 & x \\ 0 & 0 \end{bmatrix}$ is a solution for any $x \in \mathbf{C}$.

(3) $X^2 - \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} = 0$ has four solutions:

$$X = \begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 2 \end{bmatrix}.$$

It is clear that these four matrices are solutions. To see that there are no other solutions, let $X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then, as before

$$X^2 = \begin{bmatrix} a^2 + bc & b(a+d) \\ c(a+d) & bc + d^2 \end{bmatrix}.$$

If this is equal to $\begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$, then

$$0 = b(a+d) = c(a+d).$$

If $a + d \neq 0$, then $b = c = 0, a^2 = 1, d^2 = 4$, giving the four solutions listed. If $a + d = 0$, then $a = -d$ and so $a^2 = d^2$. But then, $1 = a^2 + bc = bc + d^2 = 4$, a contradiction.

(4) $X^2 - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$ has infinitely many solutions.

The same analysis as in (3) shows that

$$X = \begin{bmatrix} a & b \\ c & -a \end{bmatrix}$$

is a solution whenever $a^2 + bc = 1$. If $b = c$, these are just the matrices of reflections in $\mathbf{C}^2$.

Here is a question we will answer: For what values of l is there a quadratic equation

$$X^2 + A_1 X + A_0 = 0$$

with exactly l solutions? So far, we know that there is such an equation if $l = 0, 4$, or ∞. We have not yet seen the answer for $l = 1, 2, 3, 5, 6, 7, 8, ...$

We want to find all solutions to

$$X^2 + A_1 X + A_0 = 0. \tag{3}$$

First we need to think about how we can specify a solution X. There are several ways to do this:

(a) Write down the entries of X, e.g., $X = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$.

(b) Write down $X \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ (the first column of X) and $X \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ (the second column of X). E.g.,

$$X \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

and

$$X \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

describes the same matrix as in (a).

(c) Write down $X\mathbf{v}_1$ and $X\mathbf{v}_2$ for any basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ of $\mathbf{C}^2$. E.g.,

$$X \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and

$$X \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$$

describes the same matrix as in (a) and (b).

(d) Write down a basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for $\mathbf{C}^2$ consisting of eigenvectors for X and specify the corresponding eigenvalue for each eigenvector. (Recall that a nonzero vector $\mathbf{v}$ is an *eigenvector* for X with corresponding *eigenvalue* λ if $X\mathbf{v} = \lambda\mathbf{v}$.) Note that requiring that $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ be an eigenvector for X with corresponding eigenvalue 1 and that $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ be an eigenvector for X with corresponding eigenvalue 3 describes the same matrix as in (a), (b) and (c). Note that replacing $\mathbf{v}_i$ by $c_i\mathbf{v}_i$ where $c_1, c_2 \neq 0$ describes the same matrix. Thus, if we let

$$S = \{ \begin{bmatrix} 1 \\ u \end{bmatrix} | u \in \mathbf{C} \} \cup \{ \begin{bmatrix} 0 \\ 1 \end{bmatrix} \}$$

we may assume that the eigenvectors are chosen from S. It is important to note that not every matrix can be described in this way. The matrices that can be described this way are called *diagonalizable* matrices.

Diagonalizable solutions

We can now find all diagonalizable solutions to equation (3). Let X be such a solution and let $\mathbf{v}$ be an eigenvector for X with corresponding eigenvalue λ. Then

$$\mathbf{0} = 0\mathbf{v} = (X^2 + A_1X + A_0)\mathbf{v} = X^2\mathbf{v} + A_1X\mathbf{v} + A_0\mathbf{v}$$

$$= \lambda^2\mathbf{v} + \lambda A_1\mathbf{v} + A_0\mathbf{v} = (\lambda^2 I + \lambda A_1 + A_0)\mathbf{v}.$$

This is equivalent to requiring that

$$det(\lambda^2 I + \lambda A_1 + A_0) = 0$$

and

$$\mathbf{v} \in Null\ (\lambda^2 I + \lambda A_1 + A_0).$$

Thus λ is a root of the equation

$$det(t^2 I + tA_1 + A_0) = 0,$$

a fourth degree equation. Each of its roots is a possible eigenvalue for X and the corresponding eigenvectors may be taken from the nonzero elements of

$$S \cap Null\ (\lambda^2 I + \lambda A_1 + A_0).$$

Now the matrix

$$t^2 I + tA_1 + A_0$$

is a matrix with entries in $\mathbf{C}[t]$. A general result (rational canonical form) says that

$$P(t)(t^2 I + tA_1 + A_0)Q(t) = \begin{bmatrix} f_1(t) & 0 \\ 0 & f_2(t) \end{bmatrix}$$

where $P(t)$ and $Q(t)$ are invertible matrices over $\mathbf{C}[t]$, $det P(t) = det Q(t) = 1$ and $f_1(t), f_2(t)$ are monic polynomials in t with $f_1(t)$ dividing $f_2(t)$. (The polynomial $f_1(t)$ is the greatest common divisor of all of the entries of the matrix $t^2 I + tA_1 + A_0$.) Then $det(t^2 I + tA_1 + A_0) = f_1(t) f_2(t)$ and λ is a root of $det(t^2 I + tA_1 + A_0) = 0$ if and only if $(t - \lambda) | f_2(t)$. Furthermore, if

$$(t - \lambda) | f_1,$$

then $\lambda^2 I + \lambda A_1 + A_0 = 0$ and so

$$Null\ (\lambda^2 I + \lambda A_1 + A_0) = \mathbf{C}^2.$$

If $(t - \lambda) \not| f_1$ but $(t - \lambda) | f_2(t)$, then

$$Null\ (\lambda^2 I + \lambda A_1 + A_0) = Span\ \{Q(\lambda) \begin{bmatrix} 0 \\ 1 \end{bmatrix}\}$$

is a one-dimensional subspace.

We can now describe how to find all diagonalizable solutions of equation (3). Denote the roots of $det(t^2 I + tA_1 + A_0) = 0$ by $\lambda_i, 1 \leq i \leq m$ where $1 \leq m \leq 4$.

First suppose $f_1(t) = 1$. Then, for any $i, 1 \leq i \leq m$ we have

$$dim\ Null\ (\lambda_i^2 I + \lambda_i A_1 + A_0) = 1.$$

Therefore this space is spanned by a single vector of S. Denote this vector by $\mathbf{v}_i$. Then, for any pair $(i, j), 1 \leq i < j \leq m$ with $\mathbf{v}_i \neq \mathbf{v}_j$, there is a solution X to equation (3) such that $\mathbf{v}_i$ is an eigenvector with eigenvalue λ_i and $\mathbf{v}_j$ is an eigenvector with eigenvalue λ_j. Every diagonalizable solution to equation (3) arises in this way. Note that in this case there are at most $6 = \binom{4}{2}$ diagonalizable solutions.

Next suppose $f_1(t) = 0$ has a single root λ_1. Then, $m \leq 3$ (for $f_2(t) = 0$ has at most three roots, one of which is λ_1). In this case,

$$dim\ Null\ (\lambda_1^2 I + \lambda_1 A_1 + A_0) = 2$$

and

$$dim\ Null\ (\lambda_i^2 I + \lambda_i A_1 + A_0) = 1$$

for $2 \leq i \leq m$. Therefore, $Null\ (\lambda_i^2 I + \lambda_i A_1 + A_0)$ is spanned by a single vector of S. Denote this vector by $\mathbf{v}_i$. Now, since

$$Null\ (\lambda_1^2 I + \lambda_1 A_1 + A_0) = \mathbf{C}^2,$$

$\lambda_1 I$ is a solution. Also, if $2 \leq i \leq m$, if $\mathbf{v}_1 \in S$ and $\mathbf{v}_1 \neq \mathbf{v}_i$, there is a solution X to equation (3) such that $\mathbf{v}_1$ is an eigenvector for X with eigenvalue λ_1 and $\mathbf{v}_i$ is an eigenvector for X with eigenvalue λ_i. Here there are infinitely many choices for $\mathbf{v}_1$ and hence infinitely many solutions. In addition, if $\mathbf{v}_2 \neq \mathbf{v}_3$, there is a solution X to equation (3) such that $\mathbf{v}_2$ is an eigenvector for X with eigenvalue λ_2 and $\mathbf{v}_3$ is an eigenvector for X with eigenvalue λ_3. Every diagonizable solution to equation (3) arises in this way. Notice that in this case there are infinitely many diagonalizable solutions if and only if $m > 1$.

Finally, suppose $f_1(t) = 0$ has two roots λ_1 and λ_2. Then $f_1(t) = f_2(t)$ and λ_1 and λ_2 are the only roots of

$$det(\lambda^2 I + \lambda A_1 + A_0) = 0.$$

Furthermore,

$$dim\ Null\ (\lambda_i^2 I + \lambda_i A_1 + A_0) = 2,\ 1 \leq i \leq 2.$$

Then $\lambda_1 I, \lambda_2 I$ are solutions and for any pair $(\mathbf{v}_1, \mathbf{v}_2)$ of distinct elements of S there is a solution X to equation (3) such that $\mathbf{v}_1$ is an eigenvector for X with eigenvalue λ_1 and $\mathbf{v}_2$ is an eigenvector for X with eigenvalue λ_2. Every diagonalizable solution to equation (3) arises in this way. In this case there are infinitely many diagonalizable solutions.

This discussion goes over without serious change to the case of equations of degree n over $M_k(\mathbf{C})$. The maximum finite number of diagonalizable solutions becomes $\binom{nk}{k}$. This maximum finite number of diagonalizable solutions (and the analysis of the problem using eigenvectors and eigenvalues) is due to Fuchs and Schwarz [FS].

Examples: (1) Find all diagonalizable solutions of $X^2 - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = 0$.

The possible eigenvalues are the roots of

$$0 = det(t^2 I - \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}) = t^4.$$

Thus 0 is the only possible eigenvalue. The possible eigenvectors are the nonzero elements of

$$Null \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

These are just the nonzero multiples of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Thus we cannot find a basis for $\mathbf{C}^2$ consisting of eigenvectors for X, and so there are no diagonalizable solutions.

(2) Find all diagonalizable solutions of $X^2 = 0$.

The possible eigenvalues are the roots of

$$0 = det(t^2 I) = t^4$$

so, again, 0 is the only possible eigenvalue. The possible eigenvectors are the nonzero elements of the nullspace of the zero matrix, i.e., all nonzero vectors in $\mathbf{C}^2$. Thus we may take any basis for $\mathbf{C}^2$ and declare that both elements of the basis are eigenvectors with corresponding eigenvalues 0. This means that X is the zero matrix. Thus there is only one diagonalizable solution.

(3) Find all diagonalizable solutions of $X^2 - \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} = 0$.

The possible eigenvalues are the roots of

$$0 = det(t^2 I - \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}) = det(\begin{bmatrix} t^2 - 1 & 0 \\ 0 & t^2 - 4 \end{bmatrix})$$

$$= (t^2 - 1)(t^2 - 4) = (t - 1)(t + 1)(t - 2)(t + 2).$$

Thus the possible eigenvalues are $1, -1, 2$, and -2. If 1 or -1 is an eigenvalue, the corresponding eigenvector must be a nonzero element in $Null \begin{bmatrix} 0 & 0 \\ 0 & -3 \end{bmatrix}$, that is, a nonzero multiple of $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. If 2 or -2 is an eigenvalue, the corresponding eigenvector must be a nonzero element of $Null \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix}$, that is, a nonzero multiple of $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Thus, in determining X there are two choices (1 and -1) for the eigenvalue to assign to $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and two choices (2 and -2) for the eigenvalue to assign to $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Thus there are four diagonalizable solutions.

(4) Find all diagonalizable solutions of $X^2 - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 0$.

The possible eigenvalues are the roots of

$$0 = det(t^2 I - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}) = det(\begin{bmatrix} t^2 - 1 & 0 \\ 0 & t^2 - 1 \end{bmatrix}) = (t^2 - 1)(t^2 - 1) = (t-1)^2(t+1)^2.$$

Thus the possible eigenvalues are 1 and -1. In either case, any vector in S is a possible eigenvector. Thus we may choose $\{\mathbf{v}_1, \mathbf{v}_2\}$ to be any pair of distinct elements of S, and declare both elements to be eigenvectors for 1. This gives the identity matrix. We may also declare both elements to be eigenvectors for -1. This gives $-I$. But we may also declare $\mathbf{v}_1$ to be an eigenvector corresponding to 1 and $\mathbf{v}_2$ to be an eigenvector corresponding to -1. There are infinitely many ways to do this, so there are infinitely many solutions. (Note that if we choose an orthogonal basis for $\mathbf{C}^2$ this procedure produces the matrix of a reflection.)

Non-diagonalizable solutions

It is well-known that if a k by k matrix has k distinct eigenvalues, then the matrix is diagonalizable. Thus if a 2 by 2 matrix is not diagonalizable, it can have only one eigenvalue.

Note that if

$$X^2 + A_1 X + A_0 = 0$$

and if we set

$$Y = X + rI$$

then

$$Y^2 + (A_1 - 2rI)Y + (r^2 I - rA_1 + A_0) = 0.$$

Since λ is an eigenvalue of X if and only if $r + \lambda$ is an eigenvalue of Y, we may assume, by replacing X by $X + rI$ for an appropriate r, that 0 is an eigenvalue of X.

Since we are assuming that 0 is an eigenvalue of X we have that 0 is the only eigenvalue of X. This implies that $X^2 = 0$ (for the characteristic polynomial of X must be t^2 and by the Cayley-Hamilton Theorem, X satisfies its characteristic polynomial).

Hence it is sufficient to consider nilpotent solutions of

$$X^2 + A_1 X + A_0 = 0.$$

Since we are assuming that X is not diagonalizable, $X \neq 0$. Hence there is a linearly independent set $\{\mathbf{v}_1, \mathbf{v}_2\}$ such that

$$X\mathbf{v}_1 = \mathbf{0}$$

and

$$X\mathbf{v}_2 = \mathbf{v}_1.$$

Then

$$\mathbf{0} = 0\mathbf{v}_1 = (X^2 + A_1 X + A_0)\mathbf{v}_1$$
$$= X^2\mathbf{v}_1 + A_1 X\mathbf{v}_1 + A_0\mathbf{v}_1 = A_0\mathbf{v}_1$$

and

$$\mathbf{0} = 0\mathbf{v}_2 = (X^2 + A_1 X + A_0)\mathbf{v}_2$$
$$= X^2\mathbf{v}_2 + A_1 X\mathbf{v}_2 + A_0\mathbf{v}_2 = A_1\mathbf{v}_1 + A_0\mathbf{v}_2$$

Recall that $\mathbf{C}[t]$ denotes the algebra of polynomials in t with coefficients from the complex numbers. Let (t^2) denote the ideal in $\mathbf{C}[t]$ generated by t^2, that is, the set of all polynomials divisible by t^2. Then the quotient algebra

$$\mathbf{C}[t]/(t^2)$$

denotes the algebra of polynomials where we set $t^2 = 0$. We write

$$(\mathbf{C}[t]/(t^2))^2$$

for the set of all

$$\begin{bmatrix} g_1(t) \\ g_2(t) \end{bmatrix}$$

where $g_1(t), g_2(t) \in \mathbf{C}[t]/(t^2)$. Note that $M_2(\mathbf{C}[t])$ acts on $(\mathbf{C}[t]/(t^2))^2$ by matrix multiplication. Note also that we may write any element in $(\mathbf{C}[t]/(t^2))^2$ uniquely in the form $\mathbf{w}_1 + t\mathbf{w}_2$ where $\mathbf{w}_1, \mathbf{w}_2 \in \mathbf{C}^2$.

Now let $\mathbf{v}_1, \mathbf{v}_2$ be as above; that is,

$$0 = A_0\mathbf{v}_1$$

and

$$0 = A_1\mathbf{v}_1 + A_0\mathbf{v}_2$$

Note that this is equivalent to

$$(t^2I + tA_1 + A_0)(\mathbf{v}_1 + t\mathbf{v}_2) = 0 \tag{4}$$

in $(\mathbf{C}[t]/(t^2))^2$.

Recall that

$$P(t)(t^2I + tA_1 + A_0)Q(t)$$

$$= \begin{bmatrix} f_1(t) & 0 \\ 0 & f_2(t) \end{bmatrix})$$

where $P(t)$ and $Q(t)$ are invertible matrices over $\mathbf{C}[t]$, $detP(t) = detQ(t) = 1$ and $f_1(t), f_2(t)$ are monic polynomials in t with $f_1(t)$ dividing $f_2(t)$. The same decomposition holds over $\mathbf{C}[t]/(t^2)$.

Note that (4) has no solution unless $t^2|f_2(t)$; therefore, from now on we assume $t^2|f_2(t)$.

Note that it $t^2|f_1(t)$ then our equation is $X^2 = 0$ and the set of solutions is the set of all nilpotent matrices.

After setting $t^2 = 0$, the left-hand column of $Q(t)$ may be written as $\mathbf{p}_1 + t\mathbf{p}_2$ and the right-hand column may be written as $\mathbf{q}_1 + t\mathbf{q}_2$, where $\mathbf{p}_1, \mathbf{p}_2, \mathbf{q}_1, \mathbf{q}_2 \in \mathbf{C}^2$. Note that $\{\mathbf{p}_1, \mathbf{q}_1\}$ is linearly independent, since $\mathbf{p}_1$ and $\mathbf{q}_1$ are the columns of the invertible matrix $Q(0)$.

If $f_1(t) = t$, then the set of solutions of (4) is

$$Q(t)\begin{bmatrix} t \\ 0 \end{bmatrix} + Q(t)\begin{bmatrix} 0 \\ 1 \end{bmatrix} + Q(t)\begin{bmatrix} 0 \\ t \end{bmatrix}$$

$$= Span\ \{t\mathbf{p}_1, \mathbf{q}_1 + t\mathbf{q}_2, t\mathbf{q}_1\}.$$

Thus there is a solution X with

$$X(a\mathbf{p}_1 + \mathbf{q}_2) = \mathbf{q}_1, X(\mathbf{q}_1) = \mathbf{0}$$

whenever $\{a\mathbf{p}_1 + \mathbf{q}_2, \mathbf{q}_1\}$ is linearly independent. Since $\{\mathbf{p}_1, \mathbf{q}_1\}$ is linearly independent, there are infinitely many such solutions.

If $t \not| f_1(t)$, then (4) has solution

$$Q(t)\begin{bmatrix}0\\1\end{bmatrix} + Q(t)\begin{bmatrix}0\\t\end{bmatrix} = Span\{\mathbf{q}_1 + t\mathbf{q}_2, t\mathbf{q}_1\}.$$

If $\{\mathbf{q}_1, \mathbf{q}_2\}$ is linearly independent there is a unique nilpotent solution X with $X\mathbf{q}_2 = \mathbf{q}_1, X\mathbf{q}_1 = \mathbf{0}$. If $\{\mathbf{q}_1, \mathbf{q}_2\}$ is linearly dependent, there is no nilpotent solution.

Example: Find all nilpotent solutions of

$$X^2 - \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} = 0.$$

We write

$$\begin{bmatrix}1 & 0\\t^2 & 1\end{bmatrix}\begin{bmatrix}t^2 & -1\\0 & t^2\end{bmatrix}\begin{bmatrix}0 & 1\\-1 & t^2\end{bmatrix}$$
$$= \begin{bmatrix}1 & 0\\0 & t^4\end{bmatrix}.$$

Thus

$$\mathbf{q}_1 = \begin{bmatrix}1\\0\end{bmatrix}, \mathbf{q}_2 = \mathbf{0}$$

Since the set

$$\{\mathbf{0}, \begin{bmatrix}1\\0\end{bmatrix}\}$$

is linearly dependent, there is no nilpotent solution.

Example: Find all nilpotent solutions of

$$X^2 + \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}X + \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} = 0.$$

Here

$$\begin{bmatrix}1 & 0\\-t^2+t & 1\end{bmatrix}\begin{bmatrix}t^2+t & 1\\0 & t^2-t\end{bmatrix}\begin{bmatrix}0 & -1\\1 & t^2+t\end{bmatrix}$$
$$= \begin{bmatrix}1 & 0\\0 & t^4-t^2\end{bmatrix}.$$

Thus

$$\mathbf{q}_1 = \begin{bmatrix}-1\\0\end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix}0\\1\end{bmatrix}$$

and so there is a unique nilpotent solution.

How many solutions can there be?

We can now answer our earlier question: How many solutions can the equation

$$X^2 + A_1 X + A_0 = 0$$

have?

We have seen that if this equation has a finite number of solutions, then the number of diagonalizable solutions is at most $\binom{m}{2}$ where m is the number of distinct roots of

$$det(t^2 I + tA_1 + A_0) = 0. \tag{5}$$

We have also seen that the equation can have a nilpotent solution if and only if 0 is a root of (5) with multiplicity greater than 1. Furthermore, the number of nilpotent solutions is $0, 1$ or ∞.

Therefore, if (5) has four distinct roots and the number of solutions is finite, this number is at most 6. If (5) has a single repeated root (which we may assume to be 0) and the number of solutions is finite, there is at most 1 non-diagonalizable solution and $3 = \binom{3}{2}$ diagonalizable solutions. Finally, if (5) has two repeated roots and the number of solutions is finite, there are at most two non-diagonalizable solutions (one corresponding to each root of (5)). Also, $m = 2$, so there is at most $1 = \binom{2}{2}$ diagonalizable solution. Thus, in any case, the maximum possible finite number of solutions is 6.

The following examples show that for each $l = 0, 1, 2, 3, 4, 5, 6, \infty$ there is an equation with exactly l solutions. Each example may be analyzed using the methods above.

Examples:

(0) $X^2 + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = 0$ has no solutions.

(1a) $X^2 + \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} X + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = 0$ has a unique (nilpotent) solution.

We have

$$det \begin{bmatrix} t^2 + t & 1 \\ 0 & t^2 \end{bmatrix} = t^3(t+1),$$

and so the possible eigenvalues for a solution X are 0 and -1. The possible eigenvectors corresponding to 0 are the nonzero elements in the null space of $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix}1\\0\end{bmatrix}$. The possible eigenvectors corresponding to -1 are the nonzero elements in the null space of $\begin{bmatrix}0&1\\0&1\end{bmatrix}$. The only element of S in this space is $\begin{bmatrix}1\\0\end{bmatrix}$. Therefore there is no diagonalizable solution.

Now $\begin{bmatrix}1&0\\-t^2&1\end{bmatrix}\begin{bmatrix}t^2+t&1\\0&t^2\end{bmatrix}\begin{bmatrix}0&-1\\1&t^2+t\end{bmatrix}=\begin{bmatrix}1&0\\0&t^4+t^3\end{bmatrix}$. Thus

$$\mathbf{q}_1=\begin{bmatrix}-1\\0\end{bmatrix},\mathbf{q}_2=\begin{bmatrix}0\\1\end{bmatrix}$$

and so there is a unique nilpotent solution $\begin{bmatrix}0&-1\\0&0\end{bmatrix}$.

(1b) $X^2+\begin{bmatrix}2&0\\0&0\end{bmatrix}X+\begin{bmatrix}1&0\\0&0\end{bmatrix}=0$ has a unique (diagonalizable) solution.

We have

$$det\begin{bmatrix}t^2+2t+1&0\\0&t^2\end{bmatrix}=t^2(t+1)^2,$$

and so the possible eigenvalues for a solution X are 0 and -1. The possible eigenvectors corresponding to 0 are the nonzero elements in the null space of $\begin{bmatrix}1&0\\0&0\end{bmatrix}$. The only element of S in this space is $\begin{bmatrix}0\\1\end{bmatrix}$.The possible eigenvectors corresponding to -1 are the nonzero elements in the null space of $\begin{bmatrix}0&0\\0&1\end{bmatrix}$. The only element of S in this space is $\begin{bmatrix}1\\0\end{bmatrix}$. Thus there is a unique diagonalizable solution $\begin{bmatrix}-1&0\\0&0\end{bmatrix}$.

Now

$$\begin{bmatrix}1&1\\-(2t+3)t^2&1-(2t+3)t^2\end{bmatrix}\begin{bmatrix}(t+1)^2&0\\0&t^2\end{bmatrix}\begin{bmatrix}1-2t&-t^2\\2t+3&(t+1)^2\end{bmatrix}$$

$$=\begin{bmatrix}1&0\\0&t^2(t+1)^2\end{bmatrix}.$$

Thus

$$\mathbf{q}_1=\begin{bmatrix}0\\1\end{bmatrix},\mathbf{q}_2=\begin{bmatrix}0\\2\end{bmatrix}$$

and so there is no nilpotent solution.

A non-diagonalizable solution X of this equation with eigenvalue -1 of multiplicity 2 would correspond (upon replacing X by $X+I$) to a nilpotent solution of the equation $X^2 - \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} X + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$. Since

$$\begin{bmatrix} 1 & 1 \\ 1+(2t-3)t^2 & (2t-3)t^2 \end{bmatrix} \begin{bmatrix} t^2 & 0 \\ 0 & (t-1)^2 \end{bmatrix} \begin{bmatrix} -2t+3 & (t-1)^2 \\ 1+2t & -t^2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & t^2(t-1)^2 \end{bmatrix}.$$

Thus, in this case,

$$\mathbf{q}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

and so there is no nilpotent solution. Therefore there is no non-diagonalizable solution with eigenvalue -1 of the original equation.

(2) $X^2 + X + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = 0$ has two solutions, neither of which is diagonalizable.

We have

$$det \begin{bmatrix} t^2+t & 1 \\ 0 & t^2+t \end{bmatrix} = t^2(t+1)^2,$$

and so the possible eigenvalues for a solution X are 0 and -1. The possible eigenvectors corresponding to 0 are the nonzero elements in the null space of $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. The possible eigenvectors corresponding to -1 are the nonzero elements in the null space of $\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. Thus there are no diagonalizable solutions.

Now

$$\begin{bmatrix} 1 & 0 \\ -t^2-t & 1 \end{bmatrix} \begin{bmatrix} t^2+t & 1 \\ 0 & t^2+t \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & t^2+t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & t^2(t+1)^2 \end{bmatrix}.$$

Thus

$$\mathbf{q}_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

and so there is a unique nilpotent solution $\begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$.

A non-diagonalizable solution X of this equation with eigenvalue -1 of multiplicity 2 would correspond (upon replacing X by $X+I$) to a nilpotent solution of the equation $X^2 - X + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. Since

$$\begin{bmatrix} 1 & 0 \\ -t^2+t & 1 \end{bmatrix} \begin{bmatrix} t^2-t & 1 \\ 0 & t^2-t \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & t^2-t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & t^2(t-1)^2 \end{bmatrix}.$$

Thus, in this case,

$$\mathbf{q}_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

and so there is unique nilpotent solution $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. This corresponds to the solution

$$\begin{bmatrix} -1 & 1 \\ 0 & -1 \end{bmatrix}$$

of the original equation. This is the unique non-diagonalizable solution with eigenvalue -1.

(3) $X^2 + \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} X + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = 0$ has three solutions, one nilpotent and two diagonalizable.

We have

$$det \begin{bmatrix} t^2+t & 1-t \\ 0 & t^2-t \end{bmatrix} = t^2(t+1)(t-1),$$

and so the possible eigenvalues for a solution X are $0, 1, -1$. The possible eigenvectors corresponding to 0 are the nonzero elements in the null space of $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. The possible eigenvectors corresponding to 1 are the nonzero elements in the null space of $\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. The possible eigenvectors corresponding to -1 are the nonzero elements in the null space of $\begin{bmatrix} 0 & 2 \\ 0 & 2 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.Thus there are two diagonalizable solutions: $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$.

Now

$$\begin{bmatrix} 1 & 0 \\ -(t^3+t^2-2t)/2 & 1 \end{bmatrix} \begin{bmatrix} t^2+t & 1-t \\ 0 & t^2-t \end{bmatrix} \begin{bmatrix} \frac{1}{2} & t-1 \\ 1+\frac{t}{2} & t^2+t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & t^2(t+1)(t-1) \end{bmatrix}.$$

Thus,

$$\mathbf{q}_1 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \mathbf{q}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

and so there is a unique nilpotent solution $\begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix}$.

(4) $X^2 - \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} = 0$ has four (diagonalizable) solutions.

We have

$$det \begin{bmatrix} t^2-t & 0 \\ 0 & t^2-4t \end{bmatrix} = (t-1)(t+1)(t-2)(t+2).$$

Since this factors into distinct linear factors, all solutions are diagonalizable. We have found these solutions above.

(5) $X^2 + \begin{bmatrix} -3 & -2 \\ 0 & 1 \end{bmatrix} X + \begin{bmatrix} 2 & -2 \\ 0 & 0 \end{bmatrix} = 0$ has five (diagonalizable) solutions.

We have

$$det \begin{bmatrix} t^2-3t+2 & -2t-2 \\ 0 & t^2+t \end{bmatrix} = t(t-1)(t+1)(t-2),$$

and so the possible eigenvalues for a solution X are $0, 1, -1, 2$. The possible eigenvectors corresponding to 0 are the nonzero elements in the null space of $\begin{bmatrix} 2 & -2 \\ 0 & 0 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The possible eigenvectors corresponding to 1 are the nonzero elements in the null space of $\begin{bmatrix} 0 & -4 \\ 0 & 2 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. The possible eigenvectors corresponding to -1 are the nonzero elements in the null space of $\begin{bmatrix} 6 & 0 \\ 0 & 0 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. The possible eigenvectors corresponding to 2 are the nonzero elements in the null space of $\begin{bmatrix} 0 & -6 \\ 0 & 6 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The 5 diagonalizable solutions correspond to the pairs of eigenvalues $(0,1), (0,-1), (-1,1), (-1,2), (1,2)$.

(6) $X^2 - \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} X + \begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix} = 0$ has six (diagonalizable) solutions.

We have

$$det \begin{bmatrix} t^2 - t - 2 & t+1 \\ -2t+1 & t^2 - t - 2 \end{bmatrix} = t(t-1)(t+1)(t-2),$$

and so the possible eigenvalues for a solution X are $0, 1, -1, 2$. The possible eigenvectors corresponding to 0 are the nonzero elements in the null space of $\begin{bmatrix} -2 & 1 \\ 4 & -2 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$. The possible eigenvectors corresponding to 1 are the nonzero elements in the null space of $\begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. The possible eigenvectors corresponding to -1 are the nonzero elements in the null space of $\begin{bmatrix} 0 & 0 \\ 6 & 0 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. The possible eigenvectors corresponding to 2 are the nonzero elements in the null space of $\begin{bmatrix} 0 & 3 \\ 0 & 0 \end{bmatrix}$. The only element of S in this space is $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$. There is a diagonalizable solution corresponding to each pair of distinct possible eigenvalues. Thus the pair $(0,1)$ corresponds to $\begin{bmatrix} 2 & -1 \\ 2 & -1 \end{bmatrix}$; the pair $(0,-1)$ corresponds to $\begin{bmatrix} 0 & 0 \\ 2 & -1 \end{bmatrix}$; the pair $(1,-1)$ corresponds to $\begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}$; the pair $(-1,2)$ corresponds to $\begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$; the pair $(1,2)$ corresponds to $\begin{bmatrix} 2 & -1 \\ 0 & 1 \end{bmatrix}$; and the pair $(0,2)$ corresponds to $\begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}$.

(∞) $X^2 = 0$ has infinitely many solutions.

References

[FS] D. Fuchs and A. Schwarz, *A Matix Vieta Theorem*, E. B. Dynkin Seminar, Amer. Math. Soc. Transl. Ser 2 **169** (1996).

[GHR] I. Gohberg, P. Lancaster, and L. Rodman, *Matrix polynomials*, Academic Press, 1982.

[GR1] I. Gelfand and V. Retakh, *Noncommutative Vieta Theorem and Symmetric Functions*, Gelfand Mathematical Seminars 1993-95, Birkhauser (1996).
[GR2] I. Gelfand and V. Retakh, *Quasideterminants, I*, Selecta Math **3** (1997), no. 4.
[LR] P. Lancaster and L. Rodman, *Algebraic Riccati Equations*, Clarendon Press, 1995.

Department of Mathematics, Rutgers University, Piscataway, NJ 08854-8019
E-mail address:
`rwilson@math.rutgers.edu`