Lecture 4: calculus of equations

From the last lecture, it follows that not much is lost if we consider only $\mathfrak{C}^\infty$ solutions of $R(\frac{d}{dt})w = 0$. In fact, every $\mathfrak{L}_1^{\mathrm{loc}}$ solution is arbitrarily approximated by $\mathfrak{C}^\infty$ solutions. This gives us confidence to make the following (pretty **bold**) decision:
From this point onwards, only $\mathfrak{C}^\infty$ solutions of $R(\frac{d}{dt})w = 0$ will be considered, or, in other words, by $\mathfrak{B} = \ker\ R(\frac{d}{dt})$, we mean

$$\mathfrak{B} = \left\{ w \in \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^q) \mid R(\frac{d}{dt})w = 0 \right\} = \ker\ R(\frac{d}{dt}),$$

unless otherwise specified.

In this lecture, we intend to study some algebraic properties of polynomials and polynomial matrices, in general, required for the study of behaviors of systems. We have been dealing with linear time invariant behaviors that are represented as kernels of polynomial matrices:

$$\mathfrak{B} = \left\{ w \in \mathfrak{C}^\infty(\mathbb{R}, \mathbb{R}^q) \mid R(\frac{d}{dt})w = 0 \right\} = \ker\ R(\frac{d}{dt}),$$

where $R(\xi) \in \mathbb{R}^{g\times q}[\xi]$. It is interesting to note that, the same dynamical system can be represented by a different polynomial matrix altogether. That is, a different matrix $\widetilde{R}(\xi) \in \mathbb{R}^{g'\times q}[\xi]$ might be there, which is such that $\mathfrak{B} = \ker\ \widetilde{R}(\frac{d}{dt})$, too. This leads to the notion of *equivalent kernel representations.*

Definition 1. (equivalent systems of differential equations) *Let $R_1(\frac{d}{dt})$ and $R_2(\frac{d}{dt})$, $R_1(\xi) \in \mathbb{R}^{g\times q}[\xi]$ and $R_2(\xi) \in \mathbb{R}^{g'\times q}[\xi]$, be two kernel representation matrices. The differential equations*

$$R_1(\frac{d}{dt})w = 0 \quad \textit{and} \quad R_2(\frac{d}{dt})w = 0$$

are said to be equivalent *if they define the same dynamical system. In other words, equivalence means that w is a weak solution of $R_1(\frac{d}{dt})w = 0$* if and only if *it is also a weak solution of $R_2(\frac{d}{dt})w = 0$.*

Example 1. Let a dynamical system having two variables be described by two differential equations as:

$$\ddot{w}_1 + \ddot{w}_2 + w_1 - 2\dot{w}_2 = 0$$
$$\dddot{w}_2 + 2w_2 + 2\dot{w}_1 + 3w_1 = 0$$

The kernel representation matrix is given by:

$$R(\xi) = \begin{bmatrix} \xi^2 + 1 & \xi^2 - 2\xi \\ 2\xi + 3 & \xi^3 + 2 \end{bmatrix} \tag{1}$$

Let $r_1(\xi), r_2(\xi) \in \mathbb{R}^{1\times 2}[\xi]$ represent the two rows of the matrix $R(\xi)$. Individually $r_1(\frac{d}{dt})w = 0$ and $r_2(\frac{d}{dt})w = 0$ are the two equations that w needs to satisfy in order to qualify as an element of $\mathfrak{B} = \ker\ R(\frac{d}{dt})$. Now suppose $w \in \mathfrak{B}$ (remember that $w \in \mathfrak{C}^\infty$) be arbitrary. This means $r_i(\frac{d}{dt})w = 0$ for $i = 1, 2$. It clearly follows that if we form a new polynomial vector $r(\xi)$ as

$$r(\xi) := p(\xi)r_1(\xi) + q(\xi)r_2(\xi),$$

where $p, q \in \mathbb{R}[\xi]$ are arbitrary polynomials, then $w \in \mathfrak{B}$ should imply

$$r(\frac{d}{dt})w = 0.$$

This gives us a recipe to create newer and newer equations from a given system of equations. However, this also brings us face to face with the question: when can the original system of equations be replaced by the new equations that are generated by taking polynomial linear combinations of the original equations? In this particular example, suppose we create a new equation $\widetilde{r}_2(\xi) := p(\xi)r_1(\xi) + r_2(\xi)$, and then consider the following system of equations

$$\begin{bmatrix} r_1(\frac{d}{dt}) \\ \widetilde{r}_2(\frac{d}{dt}) \end{bmatrix} w = 0.$$

It follows that the above system of equations and the original system of equations lead to exactly same set of solutions! The reason is as follows: it is clear that if w satisfies $r_1(\frac{d}{dt})w = 0$ and $r_2(\frac{d}{dt})w = 0$ then w also satisfies $\widetilde{r}_2(\frac{d}{dt})w = 0$. Therefore, every solution of the original system is also a solution of the new system. To show the converse, let w be a solution of $r_1(\frac{d}{dt})w = 0$ and $\widetilde{r}_2(\frac{d}{dt})w = 0$. It then follows

$$\begin{aligned} \widetilde{r}_2(\frac{d}{dt})w = p(\frac{d}{dt})r_1(\frac{d}{dt})w + r_2(\frac{d}{dt})w = 0 \\ \Rightarrow r_2(\frac{d}{dt})w = 0 \ \text{ since } w \text{ satifies } r_1(\frac{d}{dt})w = 0. \end{aligned}$$

Thus, every solution of the new system is also a solution of the original system.

□

Note that the way, in this case, the new system of equations was formed is quite similar to elementary row operations performed on A by premultiplying A by an elementary row operation matrix, which does not change the solution set of $Ax = 0$. However, compared to the constant matrix with real solutions case, there are some restrictions for the polynomial case on the row operation matrices, which we will discuss shortly. In the above example, the elementary row operation matrix, say $E(\xi) \in \mathbb{R}^{2\times 2}[\xi]$, that is used to pre-multiply the original kernel representation matrix $R(\xi)$ to get the new matrix $\widetilde{R}(\xi) = E(\xi)R(\xi)$ is given by

$$E(\xi) = \begin{bmatrix} 1 & 0 \\ p(\xi) & 1 \end{bmatrix}.$$

The restriction that $E(\xi)$ must follow is that it should be *unimodular*. The reason for this restriction will soon be evident.

Definition 2. (Unimodular Matrix) *Consider $U(\xi) \in \mathbb{R}^{g\times g}[\xi]$. Then $U(\xi)$ is said to be a* unimodular *polynomial matrix if there exists a polynomial matrix $V(\xi) \in \mathbb{R}^{g\times g}[\xi]$ such that $V(\xi)U(\xi) = I$.*

A unimodular matrix is a special square polynomial matrix in the sense that its inverse also is a polynomial matrix. Note that for a polynomial matrix to have an inverse that also is a polynomial matrix it is not enough that the determinant is non-zero. For example, consider the matrix $F(\xi) = \begin{bmatrix} 1 & 1 \\ 0 & \xi \end{bmatrix}$. Now, $F(\xi)$ has determinant equal to ξ, which is non-zero, but,

$F^{-1}(\xi) = \begin{bmatrix} 1 & -\frac{1}{\xi} \\ 0 & \frac{1}{\xi} \end{bmatrix}$ is not a polynomial matrix. Theorem 1 below shows that if we have two kernel representation matrices $R(\xi)$ and $\widetilde{R}(\xi)$ related with each other by $\widetilde{R}(\xi) = U(\xi)R(\xi)$ then the two behaviours $\mathfrak{B} = \ker\ R(\frac{d}{dt})$ and $\widetilde{\mathfrak{B}} = \widetilde{R}(\frac{d}{dt})$ are equal if $U(\xi)$ is unimodular. If, on the other hand, $U(\xi)$ is *not* unimodular then the best one can say is $\widetilde{\mathfrak{B}} \supseteq \mathfrak{B}$.

Theorem 1. *Let $R(\xi) \in \mathbb{R}^{g\times q}[\xi]$ and $U(\xi) \in \mathbb{R}^{g\times g}[\xi]$. Define $\widetilde{R}(\xi) := U(\xi)R(\xi)$. Further, define $\mathfrak{B} = \ker\ R(\frac{d}{dt})$ and $\widetilde{\mathfrak{B}} = \widetilde{R}(\frac{d}{dt})$. Then,*

1. $\mathfrak{B} \subseteq \widetilde{\mathfrak{B}}$.

2. *If, in addition, $U^{-1}(\xi)$ exists and if $U^{-1}(\xi) \in \mathbb{R}^{g\times g}[\xi]$ (that is, $U(\xi)$ is unimodular), then $\mathfrak{B} = \widetilde{\mathfrak{B}}$.*

Proof: 1. Suppose $w \in \mathfrak{B}$, that is, $R(\frac{d}{dt})w = 0$. Then, $\widetilde{R}(\frac{d}{dt})w = U(\frac{d}{dt})(R(\frac{d}{dt})w) = 0 = U(\frac{d}{dt})0 = 0$. Thus, $w \in \widetilde{\mathfrak{B}}$, too. Therefore, $\mathfrak{B} \subseteq \widetilde{\mathfrak{B}}$.

2. If $U(\xi)$ is unimodular, that is, $U^{-1}(\xi)$ exists and $U^{-1}(\xi) \in \mathbb{R}^{g\times g}[\xi]$, then we have $R(\xi) = U^{-1}(\xi)\widetilde{R}(\xi)$. It then follows from Part 1. of this theorem that $\mathfrak{B} \supseteq \widetilde{\mathfrak{B}}$. Hence, we have $\mathfrak{B} = \widetilde{\mathfrak{B}}$ if $U(\xi)$ is unimodular. $\square$

Remark 1. It is quite curious that pre-multiplication matrices should be unimodular, that is, they should have inverses that themselves are polynomial matrices. It is perfectly justified to ask: *what happens if the pre-multiplication matrix $U(\xi)$ does not have a polynomial inverse?* To resolve this issue, first note that $U(\xi)$ cannot afford to have zero determinant. This is because in that case $\widetilde{R}(\xi) = U(\xi)R(\xi)$ contains strictly less number of independent equations, and therefore, $\widetilde{\mathfrak{B}} = \ker\ \widetilde{R}(\frac{d}{dt})$ must then be strictly bigger than $\mathfrak{B} = \ker\ R(\frac{d}{dt})$.

However, as we have seen in Theorem 1 above, $U(\xi)$ having just a nonzero determinant is not enough; that determinant of $U(\xi)$ must be a **nonzero real number** (and not a polynomial) in order for $\widetilde{R}(\xi)$ to be equivalent to $R(\xi)$. Now, why is that so? What is wrong with determinant being a non-constant polynomial? As an exercise one can prove that $U(\xi)$ has an inverse that also is a polynomial matrix if and only if $\det U(\xi) \in \mathbb{R} \setminus \{0\}$. So, if $\det U(\xi)$ is a non-constant polynomial then $U^{-1}(\xi)$ will contain entries that are ratios of polynomials (called *rational functions*). Now, remember that ξ is a place-holder for $\frac{d}{dt}$, and $\frac{d}{dt}$ isn't just a symbol – it has an operational meaning as the differential operator. This fact makes us avoid rational functions in the study of equivalent kernel representations: how does one make an operational sense to ξ^{-1}? Does ξ^{-1} mean *integration*? Well, it might be viewed as an integration, but, in that case, pre-multiplication by a rational function matrix cannot give equivalent kernel representation because the right hand side would then become non-zero. For example, consider the kernel representation

$$\mathfrak{B} = \ker\ \frac{d}{dt}.$$

The behaviour here has a nice parametric description as

$$\mathfrak{B} = \{k \mid k \in \mathbb{R}\}.$$

However, look at the kernel representation matrix $R(\xi) = \xi$. If pre-multiplication by rational functions were allowed then we get another kernel representation matrix $\widetilde{R}(\xi) = \xi^{-1}\xi = 1$. Then we get $\ker\ \widetilde{R}(\xi) = \{0\}$.

Thus, $R(\frac{d}{dt})w = 0 \ \Rightarrow\ U(\frac{d}{dt})R(\frac{d}{dt})w = 0$ is true if and only if $U(\xi)$ is a polynomial matrix and is *not* a matrix with rational function entries. That is why, if $U(\xi)$ is not unimodular then, with what we have learnt so far, we cannot comment on whether $\widetilde{\mathfrak{B}} = \mathfrak{B}$ or not.

Upper-triangularization

The next aim is to solve for a behavior from the kernel representation matrix. One of the ways to get the behavior is to convert $R(\xi)$ to an upper triangular form and then solve the differential equations backwards (the last equation in the triangular form will be a differential equation in one variable). The following theorem gives the construction of such a transformation.

Theorem 2. *There exists a unimodular matrix $U(\xi) \in \mathbb{R}^{g\times g}[\xi]$ such that $U(\xi)R(\xi) = T(\xi)$ and $T_{ij}(\xi) = 0$ for $i = 1, 2, \ldots, n$, $j < i$, that is $T(\xi)$ is upper triangular.*

Proof: Choose a nonzero column of $R(\xi)$. Without loss of generality, this can be taken to be the first column. In that column select the polynomial having the least degree among the entries in the column. Let that be $p_{j1}(\xi)$ where j is the row index of the polynomial. Make the j-th row the first row by premultiplying $R(\xi)$ by an elementary row operation matrix (namely, a *permutation matrix*). The row operation matrix is unimodular and thus the behavior remains the same. Now divide the polynomials $p_{i1}(\xi)$ where $i = 2, \ldots, g$ by $p_{11}(\xi)$ using long division method yielding a remainder. We get $p_{i1}(\xi) - q_{i1}(\xi)p_{11}(\xi) = r_{i1}(\xi)$ where deg $r_{i1}(\xi) <$ deg $p_{11}(\xi)$. What we have done is represented in matrix form:

$$R_1(\xi) = U_1(\xi)R(\xi) = \begin{bmatrix} p_{11}(\xi) & p_{12}(\xi) & \ldots & p_{1q}(\xi) \\ p_{21}(\xi) & p_{22}(\xi) & \ldots & p_{2q}(\xi) \\ \vdots & & \ddots & \\ p_{g1}(\xi) & p_{g2}(\xi) & \ldots & p_{gq}(\xi) \end{bmatrix},$$

where $R(\xi)$ was the original kernel representation matrix and $R_1(\xi)$ is the matrix after the minimal degree polynomial of the first column is placed at $(1,1)$ position. After doing the long division the matrix $R_1(\xi)$ is changed to $R_2(\xi)$ as

$$R_2(\xi) = U_2(\xi)R_1(\xi) = \begin{bmatrix} p_{11}(\xi) & p_{12}(\xi) & \ldots & p_{1q}(\xi) \\ r_{21}(\xi) & p_{22}(\xi) & \ldots & p_{2q}(\xi) \\ \vdots & & \ddots & \\ r_{g1}(\xi) & p_{g2}(\xi) & \ldots & p_{gq}(\xi) \end{bmatrix},$$

where $U_2(\xi)$ is given by

$$U_2(\xi) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -q_{21}(\xi) & 1 & 0 & \cdots & 0 \\ -q_{31}(\xi) & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -q_{g1}(\xi) & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Note that det $U_2(\xi) = 1$, therefore, $U_2(\xi)$ is unimodular. Thus the behavior of ker $R_2(\frac{d}{dt})$ remains the same as that of ker $R(\frac{d}{dt})$. Now, it is important to notice that, since deg $r_{i1}(\xi) <$ deg $p_{11}(\xi)$, the process of searching for the nonzero entry of minimal degree again and interchange the first row and the row in which this entry appears, makes the degree of the $(1,1)$ entry strictly less than that of $p_{11}(\xi)$ of $R_1(\xi)$. (This is achieved by premultiplication by a permutation matrix, which is unimodular.) Now, we repeat the same division with remainder procedure, which, as we have already seen, is pre-multiplying by a suitable unimodular matrix. Every time we apply this procedure, the minimal degree decreases by at least one. Also, we can apply the procedure as long as more than one entry in the first column is nonzero. Since degrees are always nonnegative, this process stops after a finite number of steps. We have then transformed the first column into a column consisting of a nonzero element in the $(1,1)$ position and all the other elements being zero. This is represented by say, $\tilde{R}(\xi)$.

$$\tilde{R}(\xi) = U(\xi)R(\xi) = \begin{bmatrix} p_{11}(\xi) & p_{12}(\xi) & \dots & p_{1q}(\xi) \\ 0 & p_{22}(\xi) & \dots & p_{2q}(\xi) \\ \vdots & & \ddots & \\ 0 & p_{g2}(\xi) & \dots & p_{gq}(\xi) \end{bmatrix}$$

where $U(\xi)$ is the product of a finite number of elementary unimodular matrices.
This has been achieved by premultiplication by elementary unimodular matrices a finite number of times. A finite product of elementary unimodular matrices is unimodular. Using an iterative procedure the next block can be transformed having a nonzero entry at the $(2,2)$ position and zero below it. Applying the procedure for consecutive subblocks we ultimately get a upper triangular form. This completes the proof. □

Definition 3. Independence of polynomial vectors: *The polynomial vectors* $r_1(\xi), r_2(\xi), \dots, r_g(\xi) \in \mathbb{R}^{1\times q}[\xi]$ *are* dependent *if there exists* $a_i(\xi) \in \mathbb{R}[\xi], 1 \leq i \leq g$ *not all zero such that,*

$$\sum_{i=1}^{g} a_i(\xi) r_i(\xi) = 0$$

Otherwise they are independent*.*

Definition 4. Row rank and Column rank: *Let* $R(\xi) \in \mathbb{R}^{g\times q}[\xi]$*. The* row rank (column rank) *of* $R(\xi)$ *is defined as the maximal number of independent rows (columns).*