UNIT VI: THE LINEAR SYSTEMS

We study linear systems of \( n \) first-order differential equations. They are related to first-order matrix differential equations. When the corresponding matrix is constant, then the eigenvalues and the eigenfunctions of the matrix provide a useful framework to construct the general solution. The fundamental matrix is constructed as the exponential matrix.

LECTURE 25. LINEAR SYSTEMS

A linear system of \( n \) differential equations in \( n \) unknowns is

\[
\begin{align*}
y_1' &= a_{11}(t)y_1 + \cdots + a_{1n}(t)y_n + f_1(t), \\
y_2' &= a_{21}(t)y_1 + \cdots + a_{2n}(t)y_n + f_2(t), \\
&\vdots &\vdots \\
y_n' &= a_{n1}(t)y_1 + \cdots + a_{nn}(t)y_n + f_n(t).
\end{align*}
\]

(25.1)

Here and elsewhere, \( \cdot' = d/dt \). In the matrix notation,

\[
\ddot{y}' = A(t)\ddot{y} + \ddot{f}(t),
\]

(25.2)

where \( \ddot{y} = (y_1(t), \ldots, y_n(t))^T \) and \( \ddot{f}(t) = (f_1(t), \ldots, f_n(t))^T \) are vector-valued functions defined on an interval \( t \in I \) and taking values in \( \mathbb{R}^n \), and \( A(t) = (a_{ij}(t))_{i,j=1}^n \) is an \( n \times n \) matrix-valued function defined on \( I \).

A matrix-valued function \( A(t) = (a_{ij}(t)) \) is said to be continuous, bounded, or differentiable if each element \( a_{ij} \) of \( A(t) \) is continuous, bounded and differentiable, respectively. Differentiation and integration are element-wise:

\[
\frac{d}{dt}A(t) = \left( \frac{da_{ij}}{dt}(t) \right), \quad \int A(t) \, dt = \left( \int a_{ij} \, dt \right).
\]

Let us define an operator

\[
L\ddot{y} = \ddot{y}' - A\ddot{y}.
\]

(25.3)

With this notation, (25.2) is written as \( L\ddot{y} = \ddot{f}(t) \). The domain of the operator \( L \) is the space of \( n \)-dimensional vector-valued functions differentiable on \( I \).

Exercise. Show that \( L \) is linear. That is,

\[
L(c_1\ddot{y}_1(t) + c_2\ddot{y}_2(t)) = c_1L\ddot{y}_1(t) + c_2L\ddot{y}_2(t).
\]

Since \( L \) is linear, the fundamental principle of superposition and the principle of the complementary function and other results for linear operators apply.

Existence and Uniqueness result. If \( A(t) \) and \( \ddot{f}(t) \) are continuous and bounded on an interval \( I \) containing \( t_0 \), then for each \( \ddot{y}_0 \in \mathbb{R}^n \) the initial value problem

\[
\ddot{y}' = A(t)\ddot{y} + \ddot{f}(t), \quad \ddot{y}(t_0) = \ddot{y}_0
\]

Exercise
has a unique solution in \( I \).

**Working assumption.** Throughout the note and the following notes, \( A(t) \) and \( f(t) \) are assumed to be continuous and bounded on an interval \( t \in I \).

**Linear independence.** The vectors \( \vec{y}_1, \ldots, \vec{y}_n \) in \( \mathbb{R}^n \) are said to be linearly independent if
\[
c_1 \vec{y}_1 + c_2 \vec{y}_2 + \cdots + c_n \vec{y}_n = 0 \quad \text{implies} \quad c_1 = c_2 = \cdots = c_n = 0.
\]
Let \( Y = (\vec{y}_1, \ldots, \vec{y}_n) \) be an \( n \times n \) matrix whose \( j \)-th column is \( \vec{y}_j \). Then, \( \vec{y}_1, \ldots, \vec{y}_n \) are linearly independent if and only if the determinant \( |Y| \neq 0 \). In this case, moreover, these functions form a basis for the linear space \( \mathbb{R}^n \).

Similarly, the vector-valued functions \( \vec{y}_1(t), \ldots, \vec{y}_n(t) \) are said to be linearly independent on the interval \( I \) if
\[
c_1 \vec{y}_1(t) + c_2 \vec{y}_2(t) + \cdots + c_n \vec{y}_n(t) = 0 \quad \text{on} \quad I \quad \text{implies} \quad c_1 = c_2 = \cdots = c_n = 0.
\]
This definition is more restrictive than that for vectors. But, if the functions are solutions of linear systems of differential equations, then their linear independence can be characterized in terms of the determinant of a matrix-valued function, analogous to that for vectors.

A basis of solutions of
\[
(25.4) \qquad \vec{y}' = A(t)\vec{y},
\]
where \( A(t) \) is a continuous \( n \times n \) matrix-valued function on \( I \), is a collection of \( n \) solutions \( \vec{y}_1(t), \ldots, \vec{y}_n(t) \) so that every solution can be expressed uniquely as their linear combination
\[
\vec{y}(t) = c_1 \vec{y}_1(t) + \cdots + c_n \vec{y}_n(t)
\]
for constants \( c_1, c_2, \ldots, c_n \).

**Theorem 25.1.** Let \( \vec{y}_1(t), \ldots, \vec{y}_n(t) \) be \( n \) solutions of (25.4) on \( I \). Let \( Y(t) = (\vec{y}_1(t), \ldots, \vec{y}_n(t)) \) be the \( n \times n \) matrix-valued function, whose \( j \)-th column is \( \vec{y}_j(t) \).

The functions form a basis of solutions if and only if \( |Y(t_0)| \neq 0 \) at some point \( t_0 \in I \). If \( |Y(t_0)| = 0 \) at a point \( t_0 \in I \), then \( |Y(t)| = 0 \) for all \( t \in I \).

If \( |Y(t_0)| = 0 \) then the column vectors in \( Y(t_0) \) are linearly dependent. Thus, we can find constants \( c_j \) not all zero such that
\[
y(t) = c_1 \vec{y}_1(t) + \cdots + c_n \vec{y}_n(t)
\]
satisfies \( y(t_0) = 0 \). Moreover, \( y(t) \) is a solution of (25.4). Then, by uniqueness, \( y(t) = 0 \) for all \( t \in I \).

That is, the functions are linearly dependent on \( I \).

**Notations.** \( |Y(t)| = \text{det}(\vec{y}_1(t) \cdots \vec{y}_n(t)) \) is called the Wronskian of \( \vec{y}_1(t), \ldots, \vec{y}_n(t) \).

If the columns of \( Y(t) \) are linearly independent solutions of \( \vec{y}' = A(t)\vec{y} \) then \( Y(t) \) is called the fundamental matrix for \( \vec{y}' = A(t)\vec{y} \). The fundamental matrix yields the general solution, and hence the complete solution, of \( \vec{y}' = A(t)\vec{y} \). This proves the assertion.

**Plane autonomous systems.** When \( n = 2 \), then the system
\[
(25.5)
\begin{align*}
y_1' &= a_{11}(t)y_1 + a_{12}(t)y_2, \\
y_2' &= a_{21}(t)y_1 + a_{22}(t)y_2
\end{align*}
\]
is called a plane system. In the matrix notation, we write it as \( \vec{y}' = A(t)\vec{y} \), where \( \vec{y} = (y_1, y_2) \) and \( A(t) = \begin{pmatrix} a_{11}(t) & a_{12}(t) \\ a_{21}(t) & a_{22}(t) \end{pmatrix} \).
When \( a_{ij} \) are constants, then let \( \phi = y_1 \). We eliminate \( y_2 \) from the system, and 
\[
(\phi' - a_{11}\phi')' = a_{12}a_{21}\phi + a_{22}(\phi' - a_{11}\phi).
\]
Thus, we obtain the *companion or secular equation*
\begin{equation}
(25.6) \quad \phi'' - (\text{tr} \ A)\phi' + (\text{det} \ A)\phi = 0
\end{equation}
of (25.5).

**Proposition 25.2.** If \((y_1(t), y_2(t))^T\) is a solution of (25.5), then both \(y_1(t)\) and \(y_2(t)\) are solutions of the secular equation (25.6).

The proof is easy and it is left as an exercise.

Conversely, the secular equation (25.6) can be used to solve the system (25.5), by finding the roots of the corresponding characteristic polynomial
\[
\lambda^2 - (\text{tr} \ A)\lambda + \text{det} \ A,
\]
and subsequently (25.5) can be solved. In this note, we call the quadratic polynomial as the *characteristic polynomial* of \( A \).

**Example 25.3.** We consider the linear system
\[
\begin{align*}
y_1' &= 12y_1 + 5y_2, \\
y_2' &= -6y_1 + y_2,
\end{align*}
\]
\( A = \begin{pmatrix} 12 & 5 \\ -6 & 1 \end{pmatrix} \).

Its secular equations is \( \phi'' - 13\phi' + 42\phi = 0 \), and the solutions are generated by \( e^{6t} \) and \( e^{7t} \).

If \( y_1(t) = c_1 e^{6t} \), then \( 5y_2 = y_1' - 12y_1 = -6c_1 e^{6t} \). If \( y_1(t) = c_2 e^{7t} \) then \( y_2(t) = -c_2 e^{7t} \). Thus, the general solutions of the system is
\[
c_1 e^{6t} \begin{pmatrix} 5 \\ -6 \end{pmatrix} + c_2 e^{7t} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.
\]

**Exercise.** Find the general solution of
\begin{enumerate}
  \item \( A = \begin{pmatrix} 6 & 1 \\ -1 & 8 \end{pmatrix} \). \text{(Answer.} \ c_1 e^t \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^t \begin{pmatrix} \frac{t}{2} \\ -t + 1 \end{pmatrix}).
  \item \( A = \begin{pmatrix} 1 & 1 \\ -4 & 1 \end{pmatrix} \). \text{(Answer.} \ c_1 e^t \begin{pmatrix} \cos 2t \\ -2 \sin 2t \end{pmatrix} + c_2 e^t \begin{pmatrix} \sin 2t \\ 2 \cos 2t \end{pmatrix} \).
\end{enumerate}

**Review of linear algebra.** We end this note by listing notations and results of linear algebra, useful in the study of linear system of differential equations.

Let \( A = (a_{ij}) \) be an \( n \times n \) matrix. If the rows and the columns of \( A \) are interchanged then the resulting matrix is called the *transpose* of \( A \), denoted by \( A^T \). It is immediate that
\[
(A^T)^T = A, \quad (A^T)^{-1} = (A^{-1})^T, \quad (AB)^T = B^T A^T.
\]

A matrix is said to be *symmetric* if \( A = A^T \).

We use \( \lvert A\rvert \) for the determinant of \( A \). The \((i, j)\)-minor of \( A \), denoted by \( M_{ij} \), is the determinant of the \((n - 1) \times (n - 1)\) matrix obtained when the \(i\)th row and the \(j\)th column of \( A \) are deleted. Then, by definition,
\[
\lvert A\rvert = a_{11}M_{11} - a_{12}M_{12} + \cdots + (-1)^{n+1}a_{1n}M_{1n}.
\]

The cofactor of \( A \), denoted by \( \text{cof} \ A \), is a \( n \times n \) matrix \((c_{ij})_{i,j=1}^n\), where
\[
c_{ij} = (-1)^{i+j}M_{ij}.
\]
Then,
\[ |A| = a_{11}c_{11} + a_{12}c_{12} + \cdots + a_{1n}c_{1n}. \]

Finally, the adjugate of $A$, denoted by $\text{adj} A$, is the transpose of the cofactor matrix
\[ \text{adj} A = (\text{cof} A)^T, \quad (\text{adj} A)_{ij} = c_{ji}. \]

Then, follows the Laplace expansion formula for determinant
\[ |A| = \sum_{j=1}^{n} a_{ij} (\text{cof} A)_{ij} = \sum_{i=1}^{n} a_{ij} (\text{cof} A)_{ij} \]
for any $1 \leq i, j \leq n$. That is a determinant can be computed on any row or on any column.

**Example 25.4.** The adjugate of the $2 \times 2$ matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $\text{adj} A = \begin{pmatrix} -d & -b \\ -c & a \end{pmatrix}$.

We state a useful formula
\[ (25.7) \quad A(\text{adj} A) = (\text{adj} A)A = (\text{det} A)I. \]
The $(i, j)$ entry of $A(\text{adj} A)$ is in fact the inner product of the $i$-th row of $A$ and the $i$-th row of $\text{cof} A$, and thus (25.7) is simply the Laplace formula.

We define the characteristic polynomial
\[ (25.8) \quad p_A(\lambda) = \text{det}(A - \lambda I) \]
of the matrix $A$. Let
\[ p_A(\lambda) = (-1)^n \lambda^n + p_1 \lambda^{n-1} + \cdots + p_0. \]
It naturally leads to a matrix polynomial
\[ p_A(A) = (-1)^n A^n + p_1 A^{n-1} + \cdots + p_0 I. \]
An important property of the adjugate is
\[ (25.9) \quad \text{adj} A = -((-1)^n A^{n-1} + p_1 A^{n-2} + \cdots + p_1 I). \]
Formally, $\text{adj} A = \frac{p(0) - p(A)}{A}$. 

Lecture 25

4

18.034 Spring 2009