Chapter 7: Systems of Linear Differential Equations

Philip Gressman

University of Pennsylvania
### Definition

Every vector function (with values in $\mathbb{R}^n$) which is $k$-times continuously differentiable on an interval $I$ is said to belong to $C^k(I, \mathbb{R}^n)$.

### Basic Fact

The space $C^k(I, \mathbb{R}^n)$ is a vector space over the reals under pointwise addition and scalar multiplication. This vector space is infinite-dimensional.

### Important Transformations

Differentiation maps $C^k(I, \mathbb{R}^n)$ to $C^{k-1}(I, \mathbb{R}^n)$ for $k > 0$ and maps $C^\infty(I, \mathbb{R}^n)$ to itself. If $A(t)$ is a $k$-times differentiable matrix-valued function, then multiplication by $A$ on the left also maps $C^k(I, \mathbb{R}^n)$ to itself.
**Linear Independence**

**Definition**

Vector functions \( x_1, \ldots, x_\ell \) are, as always, called **linearly independent** when there are no constants \( c_1, \ldots, c_\ell \) for which

\[
c_1 x_1 + \cdots + c_\ell x_\ell = 0
\]

except \( c_1 = \cdots = c_\ell = 0 \).

**IMPORTANT:** Remember that when we say that a vector function equals zero, that means it equals the old-fashioned zero vector at every single point.

Linear **dependence** is hard: if \( x_1, \ldots, x_\ell \) are linearly independent at even a single point, then as vector functions they are linearly independent. The converse is not true: they might even be linearly dependent at every point and still be linearly independent as vector functions.
We say that time-dependent vectors $\vec{X}_1, \ldots, \vec{X}_n$ are linearly independent on an interval $I$ when the only constants $c_1, \ldots, c_n$ such that

$$c_1 \vec{X}_1(t) + c_2 \vec{X}_2(t) + \cdots + c_n \vec{X}_n(t) \equiv 0$$

on the entire interval are all zeros.

Ind: $\begin{bmatrix} 1 \\ t \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \sin t \\ \cos t \end{bmatrix}$

Dep: $\begin{bmatrix} \sin^2 t \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 + t \end{bmatrix}, \begin{bmatrix} \cos^2 t \\ t \end{bmatrix}$

Ind: $\begin{bmatrix} 0 \\ t \end{bmatrix}, \begin{bmatrix} 0 \\ t^2 \end{bmatrix}$. 
The Wronskian

Given time-dependent column vectors $\vec{X}_1, \ldots, \vec{X}_n$ each of length $n$, we form the Wronskian to be the determinant of the matrix whose columns are exactly $\vec{X}_1, \ldots, \vec{X}_n$, i.e.,

$$W(\vec{X}_1, \ldots, \vec{X}_n) = \det(\vec{X}_1, \ldots, \vec{X}_n).$$

**FACT:** If the Wronskian is nonzero even at a single point, then $\vec{X}_1, \ldots, \vec{X}_n$ must be linearly independent. In fact, they might even be linearly independent when the Wronskian is always zero, but for solutions to first-order systems of ODEs this pathology does not happen.
A first-order linear system may be written in the form

$$\frac{d}{dt} \vec{X} = A(t)\vec{X} + \vec{G}(t).$$

Here $A$ is an $n \times n$ matrix whose entries may or may not depend on $t$. $\vec{G}(t)$ is a column vector of length $n$ which is fully described in the problem itself, and $\vec{X}(t)$ is an unknown column vector of length $n$ whose entries may depend on $t$.

The general solution is a complete listing of all solution vectors.

This is the specific solution for which $X(0)$ is prescribed.
• Higher-order systems of ODEs can always be recast as a system of first-order ODEs with more unknown functions.

• Systems of ODEs can always be solved by elimination; this is, however, a labor-intensive way to do it since unknown constants will be related and you’ll have to do a lot of linear equation solving.
Homogeneous Equations

Definition

The equation $\frac{d}{dt} \vec{X} = A(t) \vec{X}$ is called **homogeneous**. If your equation is given as $\frac{d}{dt} \vec{X} = A(t) \vec{X} + \vec{G}(t)$ for some nonzero $\vec{G}(t)$, then the equation $\frac{d}{dt} \vec{X} = A(t) \vec{X}$ is called the **associated homogeneous first-order system**.

Superposition Principle

If $\vec{X}_1(t)$ and $\vec{X}_2(t)$ are solutions of the homogeneous ODE $\frac{d}{dt} \vec{X} = A(t) \vec{X}$, then $c_1 \vec{X}_1 + c_2 \vec{X}_2$ will also be a solution. The same can be said for any linear combination of any number of solutions (i.e., more than two solutions).
7.3: Theory of First-order Systems

**Theorem: Existence and Uniqueness**

The IVP \( x(t_0) = x_0, \ x' = A(t)x(t) + b(t) \) for \( x, b \) vector-valued functions of time and \( A \) a matrix-valued function of time, has a unique \( C^1 \) solution on any interval \( I \) containing \( x_0 \) when \( A \) and \( b \) are continuous.

**Consequences**

**Theorem:** When \( x(t) \) is a time-dependent vector in \( \mathbb{R}^n \), the general solution of \( x'(t) = A(t)x(t) \) on any interval is an \( n \)-dimensional vector space.

**Theorem:** Solutions \( x_1, \ldots, x_n \) are linearly independent if and only if the Wronskian is never zero.

**Theorem:** If \( x_p \) is any solution to \( x' = Ax + b \), then the general solution to this ODE is given by \( x = x_c + x_p \) where \( x_c \) ranges over all solutions of the associated homogeneous equation.
Finding the general solution of an inhomogeneous system is only slightly more difficult than solving a homogeneous one.

1. You must first find some solution $\vec{X}_p$. It is called a particular solution.

2. The general solution of the inhomogeneous system will always be of the form

$$\vec{X} = c_1 \vec{X}_1 + \cdots + c_n \vec{X}_n + \vec{X}_p$$

Where $\vec{X}_1, \ldots, \vec{X}_n$ are a fundamental set of solutions (i.e., a complete set) for the associated homogeneous system.
We consider a system of ODEs with the form

\[
\frac{d}{dt} \vec{X} = A\vec{X}
\]

where \( \vec{X} \) is a column vector of length \( n \) and \( A \) is an \( n \times n \) matrix with constant entries. We begin by looking for very simple solutions, then use the superposition principle to describe the more complicated ones.

**The Simplest Case**

An example of a very simple solution is one whose direction does not change (only the magnitude). It would be expressible in the form

\[
\vec{X}(t) = f(t)\vec{E}
\]

where \( \vec{E} \) is a constant vector and \( f \) is some unknown function of \( t \).
When you assume that a solution has some special form, it is known as an *ansatz*. It’s a completely reasonable question to ask and mathematically rigorous because you might end up learning that no such solutions exist. For us, we plug our ansatz

$$\vec{X}(t) = f(t)\vec{E}$$

into the equation and get

$$f'(t)\vec{E} = f(t)A\vec{E} \Rightarrow A\vec{E} = \frac{f'(t)}{f(t)}\vec{E}.$$ 

If the equation must be true at all times, then $\frac{f'(t)}{f(t)}$ must be constant. Call the constant $\lambda$. We arrive at the eigenvector equation...
If $\vec{E}$ is an eigenvector of $A$ with eigenvalue $\lambda$, then

$$\vec{X}(t) = Ce^{\lambda t} \vec{E}$$

solves the first-order system

$$\frac{d}{dt} \vec{X} = A \vec{X}.$$ 

Moreover, linearly independent eigenvectors give linearly independent solutions of the system.

If $A$ is $n \times n$ and has $n$ linearly independent eigenvectors $\vec{E}_1, \ldots, \vec{E}_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, then the general solution of the system will be

$$\vec{X}(t) = C_1 e^{\lambda_1 t} \vec{E}_1 + \cdots + C_n e^{\lambda_n t} \vec{E}_n.$$
Complex Eigenvalues

If $A$ is a real matrix with complex eigenvalue $\lambda = \alpha + i\beta$ and eigenvector $\vec{E} = \vec{E}_{re} + i\vec{E}_{im}$, then

$$\vec{X}(t) = e^{\alpha t + i\beta t}(\vec{E}_{re} + i\vec{E}_{im})$$

will be a solution. This can only happen if the real parts and imaginary parts are each solutions by themselves. We conclude

$$\vec{X}_{re}(t) = e^{\alpha t}(\cos \beta t)\vec{E}_{re} - e^{\alpha t}(\sin \beta t)\vec{E}_{im}$$
$$\vec{X}_{im}(t) = e^{\alpha t}(\sin \beta t)\vec{E}_{re} + e^{\alpha t}(\cos \beta t)\vec{E}_{im}$$

are linearly independent real solutions of the system of ODEs.
“Missing” Eigenvectors

If $A$ does not have $n$ eigenvectors, the ansatz gives only a partial answer and we end up missing some solutions. We fix this by making a better ansatz (with increasing complexity depending on how bad the situation is). For example:

**New Ansatz**

$$\vec{X}(t) = e^{\lambda t} \vec{E}_2 + t e^{\lambda t} \vec{E}.$$  

Plug it into $\frac{d}{dt} \vec{X} = A \vec{X}$, and we get

$$e^{\lambda t} \left( \lambda \vec{E}_2 + (1 + \lambda t) \vec{E} \right) = e^{\lambda t} \left( A \vec{E}_2 + t A \vec{E} \right).$$

We must have $A \vec{E} = \lambda \vec{E}$ and $(A - \lambda I) \vec{E}_2 = \vec{E}$.

We must take $\vec{E}$ to be an eigenvector, but $\vec{E}_2$ satisfies a different equality and is called a *generalized eigenvector*. 
“Missing” Eigenvectors in General

**General Ansatz**

\[
\vec{X}(t) = e^{\lambda t} \left[ \frac{t^n}{n!} \vec{E}_n + \cdots + t \vec{E}_1 + \vec{E}_0 \right]
\]

**Generalized Eigenvectors**

The general ansatz will solve the system when

\[
(A - \lambda I) \vec{E}_n = 0,
\]

\[
(A - \lambda I) \vec{E}_{n-1} = \vec{E}_n,
\]

\[
\vdots
\]

\[
(A - \lambda I) \vec{E}_0 = \vec{E}_1.
\]
Given the first-order system

$$\frac{d}{dt} \vec{X} = A \vec{X}$$

one useful technique you should be able to use is solution by diagonalization. Here the idea is like substitution: you assume $\vec{X} = P \vec{Y}$ for some matrix $P$ and then try to solve for $Y$ instead of $X$:

$$\frac{d}{dt} P \vec{Y} = A(P \vec{Y}) \Rightarrow \frac{d}{dt} \vec{Y} = (P^{-1}AP) \vec{Y}.$$ 

So if $A$ is diagonalizable, you can do the following:

1. Solve the system

$$\frac{d}{dt} \vec{Y} = D \vec{Y}$$

where $D$ is the diagonalization of $A$.

2. To solve the original system, simply set $\vec{X} = P \vec{Y}$. 
### Matrix exponentiation

\[ e^{At} := I + tA + \frac{t^2}{2} A^2 + \frac{t^3}{3!} A^3 + \cdots \]

### Solution of the IVP

There is exactly one solution to the IVP

\[ \frac{d}{dt} \vec{X}(t) = A \vec{X}(t), \quad \vec{X}(0) = \vec{V} \]

and it equals \( \vec{X}(t) = e^{At} \vec{V} \).

There are several tricks that you might use to carry out the infinite sum and write down a simple formula that equals \( e^{At} \).
1 Use diagonalization to find a pattern for the powers $A, A^2, A^3, A^4, \ldots$. The exponential of a diagonal matrix is simply the exponential of each of the diagonal entries.

2 Solve it like a system of equations: You can write $e^{At} = b_0(t)I + b_1(t)A + \cdots + b_{n-1}(t)A^{n-1}$ for unknown functions $b_0, \ldots, b_{n-1}$. Often you can solve for these functions using the fact that

$$e^{\lambda t} = b_0(t) + b_1(t)\lambda + \cdots + b_{n-1}(t)\lambda^{n-1}$$

for each of the eigenvalues $\lambda$ (note that you will be able to solve when you have $n$ distinct eigenvalues).

3 For $2 \times 2$: if there is only one eigenvalue and only one eigenvector, then the matrix exponential will take the form

$$e^{At} = e^{\lambda t} [I + t(A - \lambda I)].$$
A phase portrait is a simultaneous plotting of several solutions of an ODE. The axes are the coordinates of the vector and the time variable is suppressed.

Two Eigenvals. $< 0$  Two Eigenvals. $> 0$  Mixed Signs

Pictures from Paul’s Online Math Notes
Phase Portraits for Complex Eigenvalues

When eigenvalue $\lambda = \alpha + i\beta$:

- $\alpha < 0$
- $\alpha = 0$
- $\alpha > 0$

Pictures from Paul’s Online Math Notes
Phase Portraits for “Missing” Eigenvectors

$\lambda < 0$  \hspace{2cm}  $\lambda > 0$

Pictures from Paul’s Online Math Notes
Inhomogeneous Systems/Undetermined Coefficients

Just like for single inhomogeneous ODEs, one can often make an educated guess about the form of the particular solution:

\[
\frac{d}{dt} \vec{X} = \begin{bmatrix} 2 & -1 \\ 0 & 3 \end{bmatrix} \vec{X} + \begin{bmatrix} 1 + e^{-t} \\ 2 \end{bmatrix} \Rightarrow \vec{X}_p = \vec{V}_1 + \vec{V}_2 e^{-t}
\]

The structure of the method is still the same:

1. Expand the inhomogeneous terms to look like vectors times constants, exponentials, powers of \( t \), and/or sines and cosines.

2. Use the tables from undetermined coeffs and/or your intuition to guess the form of the particular solution.

3. Instead of multiplying the terms in \( \vec{X}_p \) by unknown constants, multiply by unknown vectors.

4. Try to solve for the unknown vectors. If it doesn’t work, try including more terms in your guess with higher powers of \( t \) attached.