MATH 356 LECTURE NOTES
LINEAR ODES AND SYSTEMS:
LINEAR SYSTEMS IN $\mathbb{R}^n$ AND NTH ORDER ODES

J. WONG (FALL 2019)

TOPICS COVERED

• LCC systems in $\mathbb{R}^n$
  ◦ matrix exponential, use in solving $x' = Ax$
  ◦ Easy cases: basis of eigenvectors
  ◦ Jordan decomposition theorem
  ◦ Hard case: deficient eigenspace
• $n$-th order ODEs
  ◦ Translating the result to $n$-th order ODEs

1. Constant-coefficient systems

The process used to solve LCC systems in $\mathbb{R}^2$ extends to higher dimensions. Here we solve

$$x' = Ax, \quad A \in \mathbb{R}^{n \times n}. \quad (1.1)$$

From the general theory for linear systems, the space of solutions is spanned by $n$ basis functions. In reality, this is not much more than a linear algebra problem. The method is straightforward except for the case of deficient eigenspaces.

1.1. Idea: exponentials? Let $A$ have linearly independent eigenvectors $v_1, \ldots, v_n$ and eigenvalues $\lambda_1, \ldots, \lambda_n$. Then

$$x_k(t) := e^{\lambda_k t}v_k \quad (1.2)$$

is a solution for each $k$, providing a set of $n$ linearly independent solutions

$$x_1(t), x_2(t), \ldots, x_n(t).$$

To verify these solutions are a basis, we need only check linear independence at $t_0 = 0$. Since

$$\{x_k(t_0)\} = \{v_k\}$$

the solutions are indeed LI and we are done.

Complex eigenvalues: Since $A$ has real entries, complex solutions come in pairs; if $\lambda$ is a complex eigenvalue with eigenvector $v$ then $x = e^{\lambda t}v$ and $\overline{x}$ are both solutions. Then

$$\text{Re}(x), \quad \text{Im}(x)$$

are linearly independent real solutions, just as in the real case. We need only replace complex exponential solutions with real/imaginary parts.
2. Matrix exponentials

To account for all cases, we can develop some useful theory. Throughout, the notation
\[
\text{diag}(a_1, \cdots, a_n) = \begin{bmatrix}
a_1 & 0 & \cdots & 0 \\
0 & a_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & a_n
\end{bmatrix}
\]
is used for diagonal matrices (non-zero only on the main diagonal). To start, in 1d,
\[
x' = ax, \quad x(0) = x_0 \implies x = e^{at}x_0.
\]
We may guess a similar rule for \( \mathbb{R}^n \):
\[
x' = Ax, \quad x(0) = x_0 \implies x(t) = e^{At}x_0.
\]
The idea is correct, but we need to define what \( e^A \) means for an \( n \times n \) matrix.

**Definition/Properties:** For an \( n \times n \) matrix \( A \), define the **matrix exponential**
\[
e^A := \sum_{k=0}^{\infty} \frac{1}{k!} A^k = I + A + \frac{1}{2!} A^2 + \frac{1}{3!} A^3 + \cdots
\]
That is, \( e^A \) is the power series for \( e^x \) with \( x \) replaced by \( A \). It shares some properties:
1) \( e^A \) is well defined, invertible and \( (e^A)^{-1} = e^{-A} \).
2) If \( D = \text{diag}(d_1, \cdots, d_n) \) then \( e^D = \text{diag}(e^{d_1}, \cdots, e^{d_n}) \).
3) Differentiation works as expected:
\[
\frac{d}{dt}(e^{At}) = Ae^{At}.
\]
4) If \( A \) and \( B \) commute (i.e. if \( AB = BA \)) then \( e^A e^B = e^{A+B} \).
5) If \( A = VBV^{-1} \) then \( e^A = V e^B V^{-1} \).

**Warning:** Property (4) is different from the scalar \( e^x \), where \( e^{x+y} = e^x e^y \). It is **not true** that \( e^A e^B = e^{A+B} \) when \( AB \neq BA \) (due to non-commutative multiplication!).

We prove property (5) here as an example. Observe that if \( A = VBV^{-1} \) then
\[
A^k = (VBV^{-1})^k = V B^k V^{-1}
\]
and so it follows (by manipulating the series for \( e^A \) that)
\[
e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k = \sum_{k=0}^{\infty} \frac{1}{k!} V B^k V^{-1} \quad \text{(by (2.1))}
\]
\[
= V \left( \sum_{k=0}^{\infty} \frac{1}{k!} B^k \right) V^{-1} \quad \text{(factor out \( V, V^{-1} \) on left/right)}
\]
\[
= V e^B V^{-1}. \quad \text{(by def’n of \( e^B \))}
\]
In particular, if $A$ is diagonalizable and $A = V D V^{-1}$ where $D = \text{diag}(\lambda_1, \ldots, \lambda_n)$ then
\[ e^A = V e^D V^{-1}. \]

an expression that can be computed with property (2) (only scalar exponentials needed).

Note also that if $A$ is simple enough, $e^A$ can be computed directly. In particular, if $A^m = 0$ for some $m$ then the exponential is a finite sum:
\[ e^A = I + A + \frac{1}{2!} A^2 + \cdots + \frac{1}{(m-1)!} A^{m-1}. \]

2.1. Using them to solve the LCC systems. From the differentiation property (3), it is easy to show that our original guess is correct by direct substitution:

**Theorem (matrix exponential form of the solution):** Let $A$ be an $n \times n$ (real) matrix.
The solution to the initial value problem
\[ x' = Ax, \quad x(0) = x_0 \]
is given by
\[ x(t) = e^{At} x_0. \]

If $v_1, \cdots, v_n$ is a basis for $\mathbb{R}^n$ then
\[ x_i = e^{At} v_k, \quad k = 1, \ldots, n \] (2.2)
is a basis for solutions to $x' = Ax$. This formula is concise but the matrix exponential in (2.2) must be evaluated to get a ‘practical’ formula.

It is straightforward to check that this reproduces the formula we already have when $A$ has a basis of eigenvectors $\{v_k\}$. In this case, by property (5),
\[ A = V D V^{-1} \implies e^{At} = V e^{Dt} V^{-1}. \]

Then (with $e_i$ the $i$-th standard basis vector) the solutions (2.2) are
\[ e^{At} v_k = V e^{Dt} V^{-1} v_k \\
= V e^{Dt} e_k \\
= V e^{\lambda_k t} e_k (e_k \text{ is zero except } k\text{-th component}) \\
= e^{\lambda_k t} v_k \text{ (move scalar } e^{\lambda_k t} \text{ factor to the left).} \]

Note that we cannot swap the $e^{Dt}$ with $V$; only the scalar $e^{\lambda_k t}$ (hence the extra step). This is not a good use of the matrix exponential. However, if $A$ is not diagonalizable, then it provides a useful representation.

2.2. Deficient eigenspaces: theory. We now solve
\[ x' = Ax, \quad A \in \mathbb{R}^{n \times n} \]
when $A$ is not diagonalizable (i.e. it has a deficient eigenspace). To use the theorem and the formula (2.2), we must find a basis $v_1, \cdots, v_n$ for $\mathbb{R}^n$; then the solutions are
\[ x_k(t) = e^{At} v_k. \] (2.3)
However, the $v_k$’s cannot all be eigenvectors (by assumption). To correct the problem, we need one of the major theorems in linear algebra. The bare minimum is reviewed below.

**Definition:** Let $\lambda$ be an eigenvalue of an $n \times n$ matrix $A$. A **generalized eigenvector** (GE) of rank $k$ is a vector $v$ such that

$$(A - \lambda I)^k v = 0, \quad (A - \lambda I)^{k-1} v \neq 0.$$

The **chain** of GE$s$ for an eigenvalue $v$ is the sequence generated by the iteration

$$v_1 = v, \quad v_j = (A - \lambda I)v_{j+1}, \quad j = 1, 2 \cdots$$

terminating once no more solutions can be found.

**Theorem (Jordan decomposition, simplified)** Let $\lambda$ be an eigenvalue with algebraic multiplicity $k$. For each eigenvector $v$, its chain of GE$s$ $v_1, v_2, \cdots, v_m$ has finite length $m$ and is a linearly independent set.¹

The span of all the GE’s for $\lambda$ (i.e. the chains for all the eigenvectors of $\lambda$) has dimension equal to the (algebraic) multiplicity of $\lambda$.

The set of all eigenvectors of $A$ and their chains is a basis for $\mathbb{R}^n$.

For each $\lambda$ of (alg.) multiplicity $k$, the theorem guarantees $k$ linearly independent GE’s, fixing the problem of ‘not enough eigenvectors’. Figure 1 shows some possible structures.

The theorem gives us a procedure for obtaining the desired basis for $\mathbb{R}^n$ as follows:

1) Find all the distinct eigenvalues $\lambda_1, \cdots, \lambda_M$ with algebraic multiplicities $k_1, \cdots, k_M$

2) For each $\lambda$ (with alg. mult. $k$), find all the eigenvectors. Then generate the chain of GE’s for each eigenvector using (2.4) to get $k$ GE’s.

3) The result is $k_1 + \cdots + k_M = n$ GE’s, which are a basis for $\mathbb{R}^n$.

**Figure 1.** Left: a case with distinct eigenvalues. Right: a complicated case with $n = 9$. Generalized eigenvectors are shown in red. There are three eigenspaces, two are deficient ($m = 3, m = 1$ and $m = 3$ from left to right). The nine vectors form a basis for $\mathbb{R}^9$.

¹The span of the chain forms a ‘Jordan block’ for $\lambda$. The span of all the blocks for $\lambda$ is the ‘generalized eigenspace’ $\{v : (A - \lambda I)^m v = 0\}$ where $m^*$ is the maximum length of a chain for $\lambda$. 
2.3. **Deficient eigenspaces: solving the ODEs.** We compute the solutions (2.3) using the set of generalized eigenvectors as the basis. First observe that if

\[(A - \lambda I)^m v = 0\]

the matrix exponential times \(v\) is then a **finite** sum

\[e^{(A-\lambda I)t}v = v + t(A-\lambda I)v + \frac{t^2}{2}(A-\lambda I)^2v + \cdots + \frac{t^{m-1}}{(m-1)!}(A-\lambda I)^{m-1}v\]  

(2.5)

since all other terms have a factor of \((A - \lambda I)^m\). That is,

\[(A - \lambda I)^m v = 0 \implies e^{(A-\lambda I)t}v \text{ is a poly. in } t \text{ of deg. } \leq m - 1.\]

Now let \(\lambda\) be an eigenvalue with eigenvector \(v\). Its chain \(v_1, \ldots, v_m \) (with \(v_1 = v\)) yields \(m\) linearly independent solutions

\[x_j(t) = e^{\lambda t}v_j, \quad \text{for } j = 1, \ldots, m.\]

We know (from the Jordan decomposition) that \((A - \lambda I)^m v_j = 0\) for all \(j\) so by (2.5),

\[x_j(t) = e^{\lambda t}e^{(A-\lambda I)t}v_j\]

\[= e^{\lambda t} \left( v_j + t(A - \lambda I)v_j + \frac{t^2}{2}(A - \lambda I)^2v_j + \cdots + \frac{t^{m-1}}{(m-1)!}(A - \lambda I)^{m-1}v_j \right).\]  

(2.6)

By the construction of the chain, we further know that

\[(A - \lambda I)^j v_j = 0 \text{ and } (A - \lambda I)v_j = v_{j-1}.\]

Notably, the \(j\)-th solution terminates after \(j\) terms. Written out,

\[x_1(t) = e^{\lambda t}v_1\]
\[x_2(t) = e^{\lambda t}(v_2 + tv_1)\]
\[x_3(t) = e^{\lambda t}(v_3 + tv_2 + \frac{1}{2}t^2v_1)\]
\[\vdots\]
\[x_m(t) = e^{\lambda t}(v_m + tv_{m-1} + \cdots + \frac{t^{m-1}}{(m-1)!}v_1).\]

Applying this process for each eigenvector, the Jordan decomposition guarantees \(n\) solutions in total (with \(k\) solutions per eigenvalue of alg. mult. \(k\)). Linear independence is easy to check at \(t_0 = 0\) using the theorem, so we are done.

**Important conclusion:** Each basis solution for an eigenvalue \(\lambda\) of alg. mult. \(k\) looks like

\[P(t)e^{\lambda t}, \quad P = \text{polynomial of deg. } < k.\]  

(2.7)

Because we have now described all possible cases, it follows that every solution to

\[\dot{x} = Ax\]

must be a linear combination of functions in the form (2.7). That is, solutions to the LCC first-order system must look like polynomials times (complex) exponentials (in the real case, also times sin or cos). There are no other types of functions that can appear.
Example: Consider the system $\mathbf{x}' = A\mathbf{x}$ for the matrix
\[
A = \begin{bmatrix}
0 & 1 & 0 \\
0 & 0 & 1 \\
2 & 3 & 0
\end{bmatrix}.
\]
The eigenvalues are $\lambda = -1$ (alg. mult. 2) and $\lambda = 2$ (mult. 1). For $\lambda = -1$ we have one eigenvector $\mathbf{v}_1$ and a GE $\mathbf{v}_2$ (the chain has length $m = 2$) given by $\mathbf{v}_1 = (1, -1, 1)^T$, $\mathbf{v}_2 = (0, 1, -2)^T$.

This yields two solutions
\[
\mathbf{x}_1 = e^{-t}\mathbf{v}_1, \quad \mathbf{x}_2 = e^{-t}(\mathbf{v}_2 + t\mathbf{v}_1).
\]
For $\lambda = 2$, there is one eigenvector $\mathbf{v}_3 = (1, 2, 4)$ (no deficient eigenspace, so $m = 1$), which gives the solution $\mathbf{x}_3 = e^{2t}\mathbf{v}_3$.

The general solution is then
\[
\mathbf{x} = e^{-t}[c_1\mathbf{v}_1 + c_2(\mathbf{v}_2 + t\mathbf{v}_1)] + c_3e^{2t}\mathbf{v}_3 = \begin{bmatrix}
(c_1 + tc_2)e^{-t} + c_3e^{2t} \\
(-c_1 + tc_2) + c_2e^{-t} + 2c_3e^{2t} \\
(c_1 + tc_2 - 2c_2)e^{-t} + 4c_3e^{2t}
\end{bmatrix}.
\]
This is, in fact, the system version of a third order equation (see subsection 3.1).

Example: Consider the system $\mathbf{x}' = A\mathbf{x}$ for the matrix
\[
A = \begin{bmatrix}
2 & 1 & 0 & 0 \\
0 & 2 & 0 & 0 \\
0 & 0 & 2 & 1 \\
0 & 0 & 0 & 2
\end{bmatrix}.
\]
There is a single eigenvalue $\lambda = 2$ and two eigenvectors
\[
\mathbf{v}_1 = (1, 0, 0, 0)^T, \quad \mathbf{w}_1 = (0, 0, 1, 0)^T.
\]
The chain for each has length 2, with $\mathbf{v}_2 = (0, 1, 0, 0)^T, \quad \mathbf{w}_2 = (0, 0, 0, 1)^T$.

The chains for $\mathbf{v}$ and $\mathbf{w}$ give two solutions each:
\[
\mathbf{x}_1(t) = e^{2t}\mathbf{v}_1, \quad \mathbf{x}_2(t) = e^{2t}(\mathbf{v}_2 + t\mathbf{v}_1),
\]
\[
\mathbf{x}_3(t) = e^{2t}\mathbf{w}_1, \quad \mathbf{x}_4(t) = e^{2t}(\mathbf{w}_2 + t\mathbf{w}_1)
\]
and the general solution is
\[
\mathbf{x}(t) = e^{2t}[c_1\mathbf{v}_1 + c_2(\mathbf{v}_2 + t\mathbf{v}_1) + c_3\mathbf{w}_1 + c_4(\mathbf{w}_2 + t\mathbf{w}_1)] = e^{2t}\begin{bmatrix}
c_1 + tc_2 \\
c_2 \\
c_3 + tc_4 \\
c_4
\end{bmatrix}.
\]
The easier route here is to compute $e^{At} = e^{2t}e^{(A-2I)t} = I + t(A - 2I)$ which deals with both blocks at the same time.
The method used to treat second-order ODEs as special cases of systems also applies to higher order ODEs. To transform the equation
\[ y^{(n)} + a_{n-1}(t)y^{(n-1)} + a_1(t)y' + a_0(t)y = 0 \]
into a system, set \( x_i = y^{(i-1)} \) for \( i = 1, \ldots, n \) and \( x = (x_1, \ldots, x_n) \). Then
\[
\begin{align*}
  x'_1 &= x_{i+1}, \quad i = 1, \ldots, n - 1 \\
  x'_n &= -a_0x_1 - a_1x_2 - \cdots - a_{n-1}x_n.
\end{align*}
\]
Linear independence for solutions works the same way as for second order ODEs:

**Definition (linear independence, again)** A set of solutions \( y_1, \ldots, y_k \) to the \( n \)-th order ODE
\[ y^{(n)} + a_{n-1}y^{(n-1)} + \cdots + a_0y = 0 \]
is said to be linearly independent if they are linearly independent for the associated system:
\[
\begin{bmatrix}
  y_1 \\
  y'_1 \\
  \vdots \\
  y_1^{(n-1)}
\end{bmatrix}, \quad 
\begin{bmatrix}
  y_2 \\
  y'_2 \\
  \vdots \\
  y_2^{(n-1)}
\end{bmatrix}, \ldots, 
\begin{bmatrix}
  y_k \\
  y'_k \\
  \vdots \\
  y_k^{(n-1)}
\end{bmatrix}
\]
are linearly independent vectors for all \( t \). (3.1)

By the lemma for systems, this is equivalent to linear independence of the vectors at some \( t_0 \).

### 3.1. Constant coefficients
We solve the constant coefficient problem
\[ a_ny^{(n)} + a_{n-1}y^{(n-1)} + a_1y' + a_0y = 0 \] (3.2)
where \( a \)'s are real (so complex roots come in conjugate pairs). It is easiest to obtain the characteristic polynomial \( p(\lambda) \) and the eigenvalues (roots) by just plugging in \( e^{\lambda t} \) to get
\[
L[e^{\lambda t}] = p(\lambda)e^{\lambda t}, \quad p(\lambda) = a_n\lambda^n + \cdots + a_1\lambda + a_0.
\]
The solution procedure is:

1) Find the roots of the characteristic polynomial.
2b) If seeking real solutions, keep only one of each pair of complex solutions
2) Each root \( \lambda \) of multiplicity 1 gives a solution \( e^{\lambda t} \).
3) For a repeated root \( \lambda \) of multiplicity \( k \), the basis solutions are (possibly complex)
\[ e^{\lambda t}, te^{\lambda t}, \ldots, t^{k-1}e^{\lambda t}. \]

4) Take real/imaginary parts of all complex solutions.
It will always be true that this gives a basis of solutions for (3.2). Now solutions are linearly independent if at some \( t_0 \), the vectors \( y, y', \cdots, y^{(n-1)} \) are linearly independent. Linear independence of the solutions generated by this process is guaranteed by the theory for systems.

Part (3) follows almost immediately from taking the first component in the result for the associated system.\(^2\) Note that the process is much simpler than for a system!

---

\(^2\)Each \( \lambda \) with multiplicity \( k \) must have only one eigenvector. One can also check this explicitly by constructing the matrix for the system and showing that the only eigenvector for \( \lambda \) is \((1, \lambda, \cdots, \lambda^{n-1})\).
Example: The ODE
\[ y''' - 3y' - 2y = 0 \]
has characteristic polynomial
\[ p(\lambda) = \lambda^3 - 3\lambda - 2 = (\lambda - 2)(\lambda + 1)^2. \]
Since \( \lambda = -1 \) has multiplicity 2 it yields two solutions; the three basis solutions are
\[ y_1 = e^{2t}, \quad y_2 = e^{-t}, \quad y_3 = te^{-t} \]
so the general solution is
\[ y = c_1 e^{2t} + e^{-t}(c_2 + c_3 t). \]

Example: Consider the IVP
\[ y^{(4)} - y = 0, \quad y(0) = 2, \quad y'(0) = 0, \quad y''(0) = 2, \quad y'''(0) = 0. \]
From the characteristic polynomial,
\[ p(\lambda) = \lambda^4 - 1 \implies \lambda = \pm 1, \quad \lambda = \pm i. \]
This gives four solutions, two of which are complex. Taking real/imaginary parts,
\[ y(t) = c_1 e^t + c_2 e^{-t} + c_3 \cos t + c_4 \sin t. \]
From the initial conditions we get
\[
\begin{align*}
2 &= c_1 + c_2 + c_3 \\
0 &= c_1 - c_2 + c_4 \\
0 &= c_1 + c_2 - c_3 \\
0 &= c_1 - c_2 - c_4 
\end{align*}
\]
From equations 2 and 4, we find \( c_4 = 0 \); then \( c_3 = 1/2 \) and \( c_1 = c_2 = 1/2 \); the solution is
\[ y(t) = \frac{1}{2} (e^t + e^{-t}) + \cos t. \]

Remark: A more convenient choice of basis for the exponential part is
\[ \sinh t = \frac{1}{2}(e^t - e^{-t}), \quad \cosh t = \frac{1}{2}(e^t + e^{-t}) \]
so that (since \( \sinh 0 = 0 \) and \( \cosh' 0 = \sinh 0 = 0 \))
\[ y(t) = c_1 \cosh t + c_2 \sinh t + c_3 \cos t + c_4 \sin t. \]
Now applying the initial conditions, we get
\[
\begin{align*}
2 &= c_1 + c_3 \\
0 &= c_2 + c_4 \\
0 &= c_1 - c_3 \\
0 &= c_2 - c_4 
\end{align*}
\]
so the equations come in two pairs.