NOTES ON LINEAR ODES

JONATHAN LUK

We can now use all the discussions we had on linear algebra to study linear ODEs. A lot of this material is already in the textbook. These notes are essentially the same as my exposition in lectures. As always, **this is a preliminary version: if you have any comments or corrections, please send them to me**.

1. General solutions to linear systems

Given a linear ODE

$$\begin{cases} x'(t) = Ax(t) \\ x(0) = x_0, \end{cases}$$

where A is an $(n \times n)$ complex matrix and $x_0 \in \mathbb{C}^n$. We know that the unique solution is defined for all $t \in \mathbb{R}$ and is given by

$$u(t) = e^{At} u_0.$$

Given our discussions on taking exponentials of matrices, we immediately have

Theorem 1.1. *Let A be an $(n \times n)$ complex matrix. The ℓ-th component of any solution to $x'(t) = Ax(t)$ takes the following form*

$$x_\ell(t) = \sum_{\lambda_j \text{ eigenvalue of } A} p_{\ell,j}(t) e^{\lambda_j t},$$

where for each ℓ, j, $p_{\ell,j}(t)$ is a polynomial.

Proof. Recall that every matrix is similar to a matrix in the Jordan canonical form, i.e., there exists S such that $S^{-1}AS = J$, where is in the Jordan canonical form. Moreover, as we discussed earlier, every non-zero entry of e^{Jt} takes the form $e^{\lambda_i t} \frac{t^k}{k!}$ (for some eigenvalue λ_i and some $k \in \mathbb{N}$). This then implies the theorem. □

Remark 1.2. Notice that not all functions of the above form are solutions.

Remark 1.3. Suppose the characteristic polynomial is given by $\chi_A(\lambda) = \Pi_{i=1}^m (\lambda - \lambda_i)^{\nu_i}$, where λ_i are distinct. Then the polynomials $p_{\ell,j}(t)$ have degree strictly smaller than ν_j (**Exercise**: Why?)

If A is real, we want to find an expression for general real solutions. These are in particular the solutions one obtain when given real initial data. The following corollary describes what happens in that case:

Corollary 1.4. *Let A be an $(n \times n)$ real matrix. Suppose $\mu_1, \dots, \mu_m$ are all of its real eigenvalues and $\alpha_1 \pm i\beta_1, \dots, \alpha_s \pm i\beta_s$ are all of its eigenvalues in $\mathbb{C} \setminus \mathbb{R}$. Suppose moreover that the initial $x_0 \in \mathbb{R}^n$. Then the ℓ-th component of of the solution $x(t)$ takes the following form*

$$x_\ell(t) = \sum_{i=1}^m p_{\ell,i}(t) e^{\mu_i t} + \sum_{k=1}^s e^{\alpha_k t} (q_{\ell,k}(t) \sin(\beta_k t) + r_{\ell,k}(t) \cos(\beta_k t)),$$

where $p_{\ell,i}(t)$, $q_{\ell,k}(t)$ and $r_{\ell,k}(t)$ are real polynomials.

Proof. By Theorem 1.1, there exists a complex polynomial $\tilde{p}_{\ell,j}$ such that

$$\begin{aligned} x_\ell(t) = & \sum_{\lambda_j \text{ eigenvalue of } A} \tilde{p}_{\ell,j}(t)e^{\lambda_j t} \\ = & Re(\sum_{\lambda_j \text{ eigenvalue of } A} \tilde{p}_{\ell,j}(t)e^{\lambda_j t}) \\ = & \sum_{i=1}^{m} p_{\ell,i}(t)e^{\mu_i t} + \sum_{k=1}^{s} e^{\alpha_k t}(q_{\ell,k}(t)\sin(\beta_k t) + r_{\ell,k}(t)\cos(\beta_k t)) \end{aligned}$$

for some real polynomials $p_{\ell,i}(t)$, $q_{\ell,k}(t)$ and $r_{\ell,k}(t)$. □

2. Generalized eigenspaces and linear ODEs

Proposition 2.1. *Suppose $x_0 \in \mathbb{C}^n$ is an eigenvector associated to the eigenvalue λ of A. Then*

$$e^{At}x_0 = e^{\lambda t}x_0.$$

Proof. This follows from

$$e^{At}x_0 = \sum_{k=0}^{\infty} \frac{(At)^k x_0}{k!} = \sum_{k=0}^{\infty} \frac{(\lambda t)^k x_0}{k!} = e^{\lambda t}x_0.$$

□

Proposition 2.2. *Suppose $x_0 \in \mathbb{C}^n$ is a generalized eigenvector associated to the eigenvalue λ of A and $e_1, \ldots, e_\ell$ form a basis of the generalized eigenspace, then*

$$e^{At}x_0 = e^{\lambda t}\sum_{i=1}^{\ell} p_i(t)e_i,$$

where $p_i(t)$ are polynomials.

Proof. Let $A = L + N$ be the usual decomposition. Note that by the definition of L, $L^k x_0 = \lambda^k x_0$ is in the generalized eigenspace. Hence, $e^{Lt}x_0$ is in the generalized eigenspace and moreover $e^{Lt}x_0 = e^{\lambda t}x_0$.

Therefore,

$$e^{At}x_0 = e^{Nt}e^{Lt}x_0 = e^{\lambda t}e^{Nt}x_0.$$

Now note that e^{At} is in the generalized eigenspace since $A^k x_0$ is in the generalized eigenspace for all k and any subspace of $\mathbb{C}^n$ is closed (**Exercise**: Why?). This implies that e^{Nt} is in the generalized eigenspace. Finally, note that since N is nilpotent, e^{Nt} depends on t polynomially. We thus obtain our conclusion. □

Because of the above simple propositions, we can already say a lot about the solutions to an ODE if we know the eigenvalues and the generalized eigenvectors!

3. Large time behavior of solutions to linear ODEs

Proposition 3.1. *Let A be an $(n \times n)$ complex matrix such that all eigenvalues λ_j satisfy $Re(\lambda_j) < -\alpha < 0$. Then, for any solution $x(t)$ to $x'(t) = Ax(t)$, we have*

$$e^{\alpha t}\|x(t)\| \to 0$$

as $t \to \infty$.

Proof. By Theorem 1.1, each component $x_\ell(t)$ takes the form

$$u_\ell(t) = \sum_{\lambda_j \text{ eigenvalue of } A} p_{\ell,j}(t)e^{\lambda_j t}.$$

Since $Re(\lambda_j) < -\alpha < 0$, $e^{\alpha t} \times \|(\text{RHS})\| \to 0$ for each ℓ as $t \to \infty$. The proposition hence follows. □

Although Proposition 3.1 above shows that all solutions converge to 0 as $t \to \infty$, it does not actually show that 0 is a stable equilibrium. (This is because Proposition 3.1 is in principle consistent with solutions becoming very large before converging to 0). Nevertheless, we have the following more quantitative proposition[1]:

Proposition 3.2. *Let A be an $(n \times n)$ complex matrix such that all eigenvalues λ_j satisfy $Re(\lambda_j) < -\alpha < 0$. Then there exists $\Lambda > 0$ such that for every solution $x(t)$ to $x'(t) = Ax(t)$,*

$$\|x(t)\| \leq \Lambda e^{-\alpha t} \|x(0)\|$$

for all $t \geq 0$.

Proof. Since $x(t) = e^{At}x(0)$, it suffices to prove that $\|e^{At}\|_{op} \leq \Lambda e^{-\alpha t}$. By our discussions on linear algebra, there exists an invertible matrix S such that $S^{-1}AS = J$, where J is in the Jordan canonical form. Hence, $e^{At} = Se^{Jt}S^{-1}$. Recall that every non-zero entry of e^{Jt} takes the form $e^{\lambda_i t}\frac{t^k}{k!}$ (for some eigenvalue λ_i and some $k \in \mathbb{N}$). By Problem 1 in HW3, we know that $\|e^{Jt}\|_{op} \leq (\sum_{ij} |(e^{Jt})_{ij}|^2)^{\frac{1}{2}}$ (where $(e^{Jt})_{ij}$ denotes the ij-th entry of e^{Jt}). Therefore, there exists a constant $C > 0$ such that

$$e^{\alpha t}\|e^{Jt}\|_{op} \leq C \sum_{\lambda_j \text{ eignevalue of A}} e^{(\alpha + Re(\lambda_j))t}(1 + t + \cdots + t^{n-1}),$$

which is bounded for all $t \geq 0$ since $Re(\lambda_j) < -\alpha$. Finally, since

$$\|e^{At}\|_{op} \leq \|S\|_{op}\|e^{Jt}\|_{op}\|S^{-1}\|_{op},$$

we obtain the conclusion of the proposition. □

Next, we consider the case where some eigenvalues can have positive real parts. Assume that there are no eigenvalues on the imaginary axis. We first make a definition.

Recall that for

$$p_A(\lambda) = \Pi_{i=1}^m (\lambda - \lambda_i)^{\nu_i},$$

where λ_i are distinct, $\mathbb{C}^n$ can be decomposed as

$$\mathbb{C}^n = ker((\lambda_1 I - A)^{\nu_1}) \oplus \cdots \oplus ker((\lambda_k I - A)^{\nu_k}) \oplus ker((\lambda_{k+1} I - A)^{\nu_{k+1}}) \oplus \cdots \oplus ker((\lambda_m I - A)^{\nu_m}).$$

Without loss of generality (by otherwise relabelling), assume that for some $\alpha > 0$,

$$Re(\lambda_1), \dots, Re(\lambda_\ell) < -\alpha < 0$$

and

$$Re(\lambda_{\ell+1}), \dots Re(\lambda_m) > \alpha > 0.$$

Define the polynomials $p_-(\lambda)$ and $p_+(\lambda)$ by

$$p_-(\lambda) = \Pi_{i=1}^\ell (\lambda - \lambda_i)^{\nu_i}$$

and

$$p_+(\lambda) = \Pi_{i=\ell+1}^m (\lambda - \lambda_i)^{\nu_i}.$$

Clearly, $\mathbb{C}^n = \ker(p_+(A)) \oplus \ker(p_-(A))$.

Theorem 3.3. *Suppose A is a $(n \times n)$ complex matrix. Assume that there are no eigenvalues on the imaginary axis and that every eigenvalue satisfies $|Re(\lambda_i)| > \alpha$ for some $\alpha > 0$. Then*

$$e^{\alpha t}\|e^{At}x_0\| \to 0, \quad \text{as } t \to +\infty, \text{ for } x_0 \in \ker(p_-(A)), \tag{3.1}$$

$$e^{-\alpha t}\|e^{At}x_0\| \to \infty, \quad \text{as } t \to +\infty, \text{ for } x_0 \in \ker(p_+(A)) \setminus \{0\}, \tag{3.2}$$

$$e^{\alpha t}\|e^{At}x_0\| \to \infty, \quad \text{as } t \to -\infty, \text{ for } x_0 \in \ker(p_-(A)) \setminus \{0\}, \tag{3.3}$$

and

$$e^{-\alpha t}\|e^{At}x_0\| \to 0, \quad \text{as } t \to -\infty, \text{ for } x_0 \in \ker(p_+(A)). \tag{3.4}$$

[1]Note that strictly speaking this proposition only give the boundedness of $e^{\alpha t}\|u(t)\|$ instead of its decay. Nevertheless, since the inequality $Re(\lambda_i) < -\alpha < 0$ is strict, there exists $\beta > 0$ such that $Re(\lambda_i) < -\beta < -\alpha < 0$. Hence, by applying Proposition 3.2 with β in place of α, we see that its conclusion indeed implies Proposition 3.1.

Proof. We will jsut prove (3.1) and (3.2) (since the statements (3.3) and (3.4) can be proven in very similar ways.)

Proof of (3.1)**.** Let us first consider the case where $v_i \in ker((\lambda_i I - A)^{\nu_i})$ for some $i = 1, \dots, \ell$. Recall that $A = L + N$ where L is diagonalizable, $N^n = 0$ and $LN = NL$. Moreover, we showed that $Lv_i = \lambda_i v_i$. Hence, for some $C > 0$, we have

$$e^{\alpha t}\|e^{At}v_i\| = e^{\alpha t}\|e^{Lt}e^{Nt}v_i\| = e^{\alpha t}\|e^{\lambda_i t}e^{Nt}v_i\| = e^{(\alpha+Re(\lambda_i))t}\|e^{Nt}v_i\| \leq Ce^{(\alpha+Re(\lambda_i))t}(1+t+\cdots+t^{n-1}) \to 0.$$

Since any $x_0 \in \ker(p_-(A))$ can be written as a sum of such v_i's, the desired estimate for $e^{\alpha t}\|e^{At}x_0\|$ follows from the above discussion and the triangle inequality.

Proof of (3.2)**, Step 1: A reduction.** We claim that to estimate $\|e^{At}x_0\|$ for $x_0 \in \ker(p_+(A))$, it suffices to consider the case where $x_0 \in ker((\lambda_i I - A)^{\nu_i}) \setminus \{0\}$ for some $i = \ell + 1, \dots, m$. To see this, first notice that if $x_0 \in \ker(p_+(A))$, then by Proposition 2.2 $e^{At}x_0 \in \ker(p_+(A))$ for all $t \geq 0$. Moreover, we can define a norm (**Exercise**: Check that it is a norm) $\|\cdot\|_*$ on $\ker(p_+(A))$ by

$$\|v\|_* = \sum_{i=\ell+1}^{m} \|v_i\|,$$

where

$$v = v_{\ell+1} + \cdots + v_m, \quad v_i \in ker((\lambda_i I - A)^{\nu_i}).$$

Since $\ker(p_+(A))$ is finite dimensional, $\|\cdot\|_*$ must be equivalent to $\|\cdot\|$ on $\ker(p_+(A))$, i.e., there exists $C > 0$ such that for every $v \in P_+$,

$$C^{-1}\|v\| \leq \|v\|_* \leq C\|v\|.$$

Now suppose we can show that $e^{-\alpha t}\|e^{At}v_i\| \to \infty$ whenever $v_i \in ker((\lambda_i I - A)^{\nu_i}) \setminus \{0\}$ for some $i = 1, \dots, k$. Given $x_0 \in \ker(p_+(A)) \setminus \{0\}$, $x_0 = v_{\ell+1} + \cdots + v_m$, $v_i \in ker((\lambda_i I - A)^{\nu_i})$. At least one of these v_i's is nonzero. Hence,

$$e^{-\alpha t}\|e^{At}x_0\| \geq C^{-1}e^{-\alpha t}\|e^{At}x_0\|_* \geq C^{-1}\sup_i e^{-\alpha t}\|e^{At}v_i\| \to \infty.$$

Proof of (3.2)**, Step 2.** Now let $x_0 \in ker((\lambda_i I - A)^{\nu_i}) \setminus \{0\}$ for some $i = \ell + 1, \dots, m$. We have

$$\|e^{At}x_0\| = \|e^{Lt}e^{Nt}u_0\| = \|e^{\lambda_i t}e^{Nt}u_0\| = e^{Re(\lambda_i)t}\|e^{Nt}u_0\|. \tag{3.5}$$

We claim that $\|e^{Nt}u_0\|$ is bounded below for large t. To see this, note that since $N^n = 0$, $e^{Nt} = \sum_{i=0}^{n-1} \frac{t^i N^i}{i!}$. Hence, there exists a largest $i_0 \geq 0$ such that $N^{i_0}x_0 \neq 0$. Moreover, since all the other non-zero terms have smaller powers of t, for t sufficiently large, $\|e^{Nt}x_0\| \geq \frac{1}{2}\|\frac{t^{i_0}N^{i_0}}{i_0!}x_0\| =: \gamma > 0$. Hence, by (3.5), we have that if $x_0 \in ker((\lambda_i I - A)^{\nu_i}) \setminus \{0\}$ for some $i = \ell + 1, \dots, m$,

$$e^{-\alpha t}\|e^{At}x_0\| = e^{-\alpha t + Re(\lambda_i)t}\|e^{Nt}x_0\| \geq \gamma e^{-\alpha t + Re(\lambda_i)t} \to \infty.$$

Together with part 1, this concludes the proof. □

Remark 3.4. $\ker(p_-)$ is called the *stable subspace* (or *stable manifold*, or *incoming manifold*) and $\ker(p_+)$ is called the *unstable subspace* (or *unstable manifold*, or *outgoing manifold*).

Remark 3.5. As long as $\ker(p_+) \neq \{0\}$, 0 is an unstable equilibrium. Moreover, for $x_0 = x_+ + x_-$ with $x_\pm \in \ker(p_\pm)$, $\|e^{At}x_0\| \to \infty$ as long as $x_+ \neq 0$. In particular, there exists an open and dense set $U \in \mathbb{C}^n$ such that $x_0 \in U$ implies $\|e^{At}x_0\| \to \infty$.

Remark 3.6. The discussion above assumes that A is a complex matrix. In the case where A is real, in fact one can also decompose

$$\mathbb{R}^n = \ker_{\mathbb{R}^n}(p_+(A)) \oplus \ker_{\mathbb{R}^n}(p_-(A)), \tag{3.6}$$

where $\ker_{\mathbb{R}^n}(p_\pm(A)) = \ker(p_-(A)) \cap \mathbb{R}^n$.

To see this, let us note if A is an $(n \times n)$ real matrix and $\lambda_i \in \mathbb{C} \setminus \mathbb{R}$ is an eigenvalue, then its complex conjugate $\bar{\lambda}_i$ is also an eigenvalue. (This is because if $Av = \lambda v_i$ for $v \in \mathbb{C}^n \setminus \{0\}$, then $A\bar{v}_i = \bar{\lambda}\bar{v}_i$.) Therefore $p_+(\lambda)$ and $p_-(\lambda)$ are real polynomials.

As a consequence of $p_+(\lambda)$ being real, we know that if $w \in \ker(p_+(A))$ implies $\bar{w} \in \ker(p_+(A))$. Similarly for $\ker(p_-(A))$.

Now for any $v \in \mathbb{C}^n$, we decompose $v = P_+v + P_-v$ and $\bar{v} = P_+\bar{v} + P_-\bar{v}$ where $P_\pm$ is the projection to $\ker(p_\pm(A))$. By the above discussion, we know also that $\bar{v} = \overline{P_+v} + \overline{P_-v}$ and $\overline{P_\pm v} \in \ker(p_\pm(A))$. By uniqueness of the decomposition, we have

$$\overline{P_\pm v} = P_\pm \bar{v} \tag{3.7}$$

for every $v \in \mathbb{C}^n$.

By (3.7), it follows that if $v = \bar{v}$ (i.e. if v is real), then $P_\pm v = \overline{P_\pm v}$ (i.e. $P_\pm v$ is real). This then implies (3.6).

The above discussion does not cover the case where at least one eigenvalue lies on the imaginary axis. That case is in general more complicated and will be considered in the homework. In the following two examples, we will see that even if 0 is the only eigenvalue, there can be rather different long time behavior. In particular, in Example 3.7, 0 is a stable equilibrium while in Exapmle 3.8, 0 is not a stable equilibrium.

Example 3.7. Let A be the (2×2) zero matrix. In this case, any solution to $x'(t) = Ax(t) = 0$ is constant in time. In particular, it neither grows nor decays as $t \to \infty$.

Example 3.8. In the second example, consider

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

The eigenvalues are also 0 as in the previous example. Nevertheless,

$$e^{At}x_0 = \begin{bmatrix} 1 & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} (x_0)_1 \\ (x_0)_2 \end{bmatrix} = \begin{bmatrix} (x_0)_1 + t(x_0)_2 \\ (x_0)_2 \end{bmatrix}.$$

Hence, the solution grows as $t \to \infty$ if $(x_0)_2 \neq 0$.

Finally, let us note that in both Proposition 3.1 and Example 3.7, 0 is a stable equilibrium, but they have somewhat different long time behavior in that $x(t) \to 0$ in Proposition 3.1 but $x(t) \not\to 0$ in general in Example 3.7. It is therefore useful to make the following definition:

Definition 3.9 (Asymptotic stability)**.** We say that an equilibrium solution $\bar{x}$, is *asymptotically stable* if both of the following hold:

- $\bar{x}$ is stable, and
- there exists $r > 0$ such that for any data $x_0 \in B(\bar{x}, r)$, the corresponding solution $x(t)$ satisfies

$$\|x(t) - \bar{x}\| \to 0$$

as $t \to \infty$.

Remark 3.10. Let us note that the assumption that $\bar{x}$ is stable is necessary. (**Exercise**: Why?)