Math 322 lecture for Monday, Week 3

Let $F = \mathbb{R}$ or $\mathbb{C}$, and let $M_n(F)$ denote $n \times n$ matrices with coefficients in F. The derivative of a curve $x(t) = (x_1(t), \dots, x_n(t))$ in F^n with respect to t gives the curve's tangent direction or velocity at time t:

$$\dot{x} := x'(t) := \left(\frac{dx_1}{dt}, \dots, \frac{dx_n}{dt} \right).$$

We are interested in finding x such that

$$x' = Ax$$

and satisfying some initial condition $x(0) = x_0 \in F^n$. If $n = 1$, then $A = a \in F$, and we have already seen the solution $x = x_0 e^{at} = e^{at} x_0$. It turns out that the solution in the case $n = 1$ is just a space case of the solution for $n \geq 1$:

$$x = e^{At} x_0. \tag{1}$$

Our first goal is to make sense of equation (1) (e.g., what does it mean to exponentiate a matrix?) and then prove that it is the unique solution.

Definition. A *norm* on a vector space V over F is a mapping

$$\| \ \| \colon V \to \mathbb{R}$$

satisfying

1. (positive definite) $\|v\| \geq 0$ for all $v \in V$, and $\|v\| = 0$ if and only if $v = 0$.
2. (absolute homogeneity) $\|\alpha v\| = |\alpha| \|v\|$ for all $v \in V$ and $\alpha \in F$.
3. (triangle inequality) $\|v + w\| \leq \|v\| + \|w\|$ for all $v, w \in V$.

Examples. The usual absolute value on F^n is a norm. If $F = \mathbb{R}$, we have

$$\|x\| := |x| := \sqrt{x \cdot x} = \sqrt{\textstyle\sum_j x_j^2}$$

and if $F = \mathbb{C}$, we have

$$\|x\| := |x| = \sqrt{x \cdot \bar{x}} = \sqrt{\textstyle\sum_j |x_j|^2}.$$

Note: if $x_j = a_j + b_j i$ with $a_j, b_j \in \mathbb{R}$, then

$$\|x\| = |x| = \sqrt{\textstyle\sum_j (a_j^2 + b_j^2)},$$

which is the length of $x \in \mathbb{C}^n$ thought of as a vector in $\mathbb{R}^{2n}$. As indicated above, we use the usual absolute value notation, $|x|$ for this norm.

The case $n = 1$ says the usual absolute value on F is a norm on F.

Given a norm $\| \ \|$ on a vector space V, we can define a *metric* on V (i.e., a distance function) by

$$d(v, w) := \|v - w\|.$$

The following properties of this distance function are easy to verify:

1. (positive definite) $d(v, w) \geq 0$ for all $v, w \in V$, and $d(v, w) = 0$ if and only if $v = w$.
2. (symmetry) $d(v, w) = d(w, v)$ for all $v, w \in V$.
3. (triangle inequality) $d(u, w) \leq d(u, v) + d(v, w)$ for all $u, v, w \in V$.

The following proposition implies that two norms on a vector space will define the same topology ("sense of closeness") on that space:

Proposition. Let $\| \ \|_1$ and $\| \ \|_2$ be two norms on a finite-dimensional vector space V over F. Then these norms are *equivalent* in the following sense: there exist positive real numbers a, b such that

$$a\|v\|_2 \leq \|v\|_1 \leq b\|v\|_2$$

for all $v \in V$.

Sketch of proof.

Step 1. If the displayed set of inequalities holds, say $\| \ \|_1 \sim \| \ \|_2$. Prove that $\sim$ is an equivalence relation.

Step 2. By Step 1, it suffices to prove the result when $\| \ \|_2 = | \ |$, the usual absolute value norm, discussed above, and $\| \ \|_1$ is arbitrary. There is nothing to prove if $v = 0$, since any positive constants a and b work in that case. Assume from now an that $v \neq 0$. Then, dividing through by $|v|$ and using properties of the norm, we see that $a|v| \leq \|v\|_1 \leq b|v|$ is equivalent to $a \leq \|u\|_1 \leq b$ where $u = v/|v|$ has (usual) norm $|u| = 1$.

STEP 3. Show that $v \to \|v\|_1$ is a continuous function with respect to $|\ |$. That is, given $v \in V$ and $\varepsilon > 0$, show there exists $\delta > 0$ such that if $w \in V$ and $|v - w| < \delta$, then

$$|\,\|v\|_1 - \|w\|_1\,| < \varepsilon.$$

STEP 4. Apply the *extreme value theorem*, a continuous function on a compact set (closed and bounded) achieves a minimum and a maximum value. In our case, the compact set is $\{u \in V : \|u\|_1 = 1\}$ and the minimum and maximum values are the desired constants a and b, respectively. □

Definition. The *operator norm* on the vector space $M_n(F)$ of $n \times n$ matrices with coefficients in F is given by

$$\|A\| := \max_{|x| \leq 1} |Ax|.$$

for each $A \in M_n(F)$ where $|\ |$ is the usual norm on F.

Remarks.

1. For the identity matrix, we have $\|I_n\| = 1$.

2. The real number $\|A\|$ is the most that A scales any vector:

$$\|A\| = \max_{x \neq 0} A\left(\frac{x}{|x|}\right) = \max_{x \neq 0} \frac{|Ax|}{|x|}.$$

   Thus, $|Ax| \leq \|A\||x|$ for all $x \in F^n$. A detailed proof will be given below.

3. When trying to define a norm on $M_n(F)$, it might seem more natural to just think of an $n \times n$ matrix as an element of F^{n^2} and use the usual norm on F^{n^2}. However, the norm we have just described is easier to work with and, according to the proposition given above, it is equivalent to any other norm on $M_n(F)$.

Lemma 1. For all $A, B \in M_n(F)$ and $x \in F^n$,

1. $|Ax| \leq \|A\||x|$.

2. $\|AB\| \leq \|A\|\|B\|$.

3. $\|A^k\| \leq \|A\|^k$.

Proof. For part 1, first note that the inequality holds when $x = 0$. So suppose that $x \neq 0$, and let $u = \frac{x}{|x|}$. We have that $|u| = 1$, and hence,

$$\frac{|Ax|}{|x|} = \left| A\frac{x}{|x|} \right| = |Au| \leq \max_{|y|\leq 1} |Ay| = \|A\|.$$

Multiplying through by $|x|$ gives $|Ax| \leq \|A\||x|$, as desired.

For part 2, note that for all $x \in F^n$ with $|x| \leq 1$, we have from part 1,

$$|(AB)(x)| = |A(Bx)| \leq \|A\||Bx| \leq \|A\|\|B\||x| \leq \|A\|\|B\|.$$

Therefore,

$$\|AB\| := \max_{|x|\leq 1} |(AB)(x)| \leq \|A\|\|B\|.$$

Part 3 follows from part 2. □

Definition. Let $(v_k)_{k=0,1,\ldots}$ be a sequence in a normed vector space $(V, \| \|)$. We say

$$\lim_k v_k = v$$

for some vector $v \in V$ if for all $\varepsilon > 0$ there exists $N \in \mathbb{R}$ such that

$$\|v - v_k\| < \varepsilon$$

whenever $k \geq N$. A series $\sum_{k=0}^{\infty} v_k$ converges to v if its sequence of partial sums v_0, $v_0 + v_1$, $v_0 + v_1 + v_2, \ldots$ converges to v.

Theorem. For all $A \in M_n(F)$ and $t_0 > 0$, the function $\mathbb{R} \to M_n(F)$ given by

$$t \mapsto \sum_{k\geq 0} \frac{A^k t^k}{k!}$$

converges absolutely and uniformly for $t \in [-t_0, t_0]$.

Before proving this theorem, let's review the notions of absolute and uniform convergence of series of functions. First, a series $\sum_k v_k$ in a normed vector space $(V, \| \|)$ is *absolutely convergent* if $\sum_k \|v_k\|$ converges. If a series is absolutely convergent then every rearrangement of the series will converge.

Let V and W be normed vector spaces, and let $C \subseteq W$. (For instance, we could take $W = \mathbb{R}$ and $C = [-t_0, t_0]$.) For each $n \geq 0$, let $f_n \colon W \to V$ be a function. The sequence (f_n) *converges uniformly* to $f \colon W \to V$ on C if for all $\varepsilon > 0$, there exists an $N(\varepsilon) \in \mathbb{R}$ such that for all $x \in C$,

$$\|f(x) - f_n(x)\| < \varepsilon$$

whenever $n > N(\varepsilon)$. **Note:** the word "uniform" refers to the fact that $N(\varepsilon)$ is independent of x.

The notion of uniform convergence makes sense for a series $\sum_k f_k$ since a series is just a sequence of partial sums.