2. Some basic tools and theorems

2.1 Norms

For an n-dimensional vector $\mathbf{x}$ we assign a real number $\|\mathbf{x}\|$, called the norm of $\mathbf{x}$, with the following properties:

$$\begin{gathered} \|\mathbf{x}\| \geq 0 \\ \big|\|\mathbf{x}\| - \|\mathbf{z}\|\big| \leq \|\mathbf{x} + \mathbf{z}\| \leq \|\mathbf{x}\| + \|\mathbf{z}\| \\ \|\alpha\,\mathbf{x}\| = |\alpha|\|\mathbf{x}\| \\ |\mathbf{x} \cdot \mathbf{z}| \leq \|\mathbf{x}\|\|\mathbf{z}\| \end{gathered} \tag{2. 1}$$

One possible definition of the norm is

$$\|\mathbf{x}\| = \sum_{i=1}^{n} |x_i| \tag{2. 2}$$

which leads to the following (consistent) definition of the norm for a matrix

$$\|\mathbf{A}\| = \sum_{i=1}^{n} \sum_{j=1}^{n} |A_{ij}| \tag{2. 3}$$

The latter obeys the basic properties of the norm as given above.

Another possible choice for the matrix norm is

$$\|\mathbf{A}\| = \underset{\|\mathbf{a}\| \neq 0}{\mathrm{Sup}} \frac{\|\mathbf{A}\,\mathbf{a}\|}{\|\mathbf{a}\|} = \underset{\|\mathbf{a}\| = 1}{\mathrm{Sup}} \|\mathbf{A}\,\mathbf{a}\| \tag{2. 4}$$

With this choice and with the following choice for the vector norm,

$$\|\mathbf{a}\| = \underset{1 \leq i \leq n}{\mathrm{Max}} |a_i| \tag{2. 5}$$

the matrix norm becomes

$$\|\mathbf{A}\| = \underset{1 \leq i \leq n}{\mathrm{Max}} \sum_{j=1}^{n} |A_{ij}| \tag{2. 6}$$

For a diagonalizable matrix $\mathbf{A}$, with eigenvalues denoted by λ_i, if the matrix norm is defined by (Eq. (2.4)), and the vector norm is chosen as

$$\|\mathbf{a}\| = \sqrt{\sum_{i=1}^{n} |a_i|^2} \tag{2. 7}$$

then

$$\|\mathbf{A}\| = \underset{1 \leq i \leq n}{\mathrm{Max}} |\lambda_i|$$

2.2 The O, o symbols

We shall quite often need to classify the size of a quantity, as something happens. For this we use the O and o symbols. We say that the varying quantities a and b satisfy

$$a = O(b) \tag{2. 8}$$

if a positive constant K exists such that, as a and b vary,

$$|a| \le K|b| \tag{2. 9}$$

For instance, we may be interested in the degree of accuracy of an approximation, $y(t)$, to the exact (but unknown) solution, $x(t)$ to an equation that depends on a small parameter, ε, and the extent in time over which it is valid. The statement

$$|x(t) - y(t)| = O(\varepsilon^3) \qquad t = O(1/\varepsilon) \tag{2.10}$$

means that there exist two constants, M and T such that

$$|x(t) - y(t)| \le M\varepsilon^3 \tag{2.11}$$

as long as the time t obeys

$$t \le T/\varepsilon \tag{2.12}$$

Clearly, this is of interest primarily for small ε ($|\varepsilon| \ll 1$). If instead of the inequality (2.9) an equality holds, we say that

$$a = O_S(b)$$

where the subscript S stands for "strict".

We say that

$$a = o(b) \tag{2.13}$$

if

$$\frac{|a|}{|b|} \xrightarrow[b\to 0]{} 0 \tag{2.14}$$

Similar definitions hold for vectors, in terms of their norms. For instance, we would say that $\mathbf{x}=o(\varepsilon)$ if $\|\mathbf{x}\|$ vanishes more rapidly than ε when ε tends to zero, or that $\mathbf{x}=o(1/M)$ if $\|\mathbf{x}\|$ vanishes more rapidly than $1/M$ when M tends to infinity.

2.3 The Lipschitz condition

An n-dimensional vector function $\mathbf{f}(\mathbf{x})$ defined over a domain D in R^n satisfies the Lipschitz condition in D if a positive constant L exists such that for any two points $\mathbf{x}_1, \mathbf{x}_2 \in D$

$$\|\mathbf{f}(\mathbf{x}_1) - \mathbf{f}(\mathbf{x}_2)\| \le L\|\mathbf{x}_1 - \mathbf{x}_2\| \tag{2.15}$$

This clearly requires continuity. One simple example in one dimension is a function with a bounded derivative in an interval [a,b]. The *mean value theorem* states that

$$f(x_1) - f(x_2) = f'(\xi)(x_1 - x_2) \qquad x_1 \le \xi \le x_2 \tag{2.16}$$

which yields (here the norm is simply the absolute value)

$$|f(x_1) - f(x_2)| = \underset{a \le \xi \le b}{\text{Max}} |f'(\xi)| |(x_1 - x_2)| \tag{2.17}$$

The Lipschitz condition is one of the most common conditions guaranteeing the existence and uniqueness of the solutions of ordinary differential equations.

<u>Theorem (without proof)</u> - Existence and uniqueness of solutions

The initial value problem

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, t; \varepsilon) \qquad \mathbf{x}(0) = \mathbf{x}_0$$
$$\mathbf{x} \in D \subseteq \mathrm{R}^n \qquad \mathbf{f} \in \mathrm{R}^n \tag{2.18}$$

where **f** satisfies the Lipschitz condition in D for $t_0 \le t \le t_0 + T$ and for $0 \le \varepsilon \le \varepsilon_0$ has a unique solution for times $t_0 \le t \le t_0 + \tau$ where

$$\tau = \inf\left(T, \frac{d}{M}\right)$$
$$M = \sup_{\substack{\mathbf{x} \in D \\ t_0 \le t \le t + T \\ 0 \le \varepsilon \le \varepsilon_0}} \|\mathbf{f}(\mathbf{x})\| \qquad d = \inf_{x \in \partial D} \|\mathbf{x} - \mathbf{x}_0\| \tag{2.19}$$

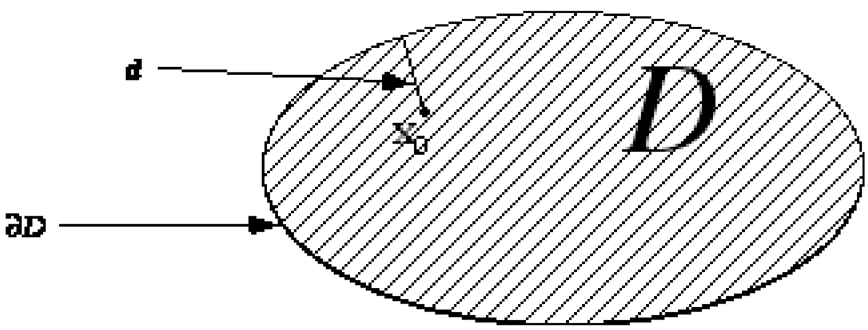
and ∂D is the boundary of D.

Note that d/M provides a rough bound for the shortest time it takes an initial condition at $t = t_0$ to propagate to the edge of the domain D. Also, the temporal extent of the existence of the solution is limited by the size of the maximum norm of **f**(**x**). For instance, if one has

$$\mathbf{f} = \varepsilon \mathbf{g}$$

with **g** of bounded norm, then

$$M = \sup\|\mathbf{f}\| = |\varepsilon| \sup\|\mathbf{g}\|$$



<u>Fig. 2.1</u>

and

$$\tau = \inf\left(T, \frac{d}{\varepsilon \sup\|\mathbf{g}\|} \right) \tag{2.20}$$

Thus, the solution may exist and be unique for times O(1/ε). This is the case, for instance, if **f**(**x**) satisfies the Lipschitz condition up to infinite times.

2.3.1 Lipschitz condition-intuitive picture in one dimension

We follow here the analysis of Arnold [6]. Consider the one dimensional case of Eq. (2.11) with $f(x)$ satisfying the Lipschitz condition. Assume that there are two *different* solutions, $x_1(t)$ and $x_2(t)$, both satisfying the same initial condition at t=0. Now define

$$y(t) \equiv x_1(t) - x_2(t)$$

which is not identically zero. However, $y(t)$ does vanish at some points (at least at t=0). Let t_1 be a point at which $y(t)$ vanishes, and assume (without loss of generality) that for some range of times $t \geq t_1$ $y(t)$ is positive and monotonically increasing. Positivity and monotonicity of y(t) yields

$$\begin{aligned} 0 \leq \frac{dy}{dt} &= \frac{d(x_1 - x_2)}{dt} = f(x_1) - f(x_2) \\ &= \left|\frac{dy}{dt}\right| = \left|f(x_1) - f(x_2)\right| \leq L\left|x_1 - x_2\right| = L|y| = Ly \end{aligned} \tag{2.21}$$

Choose some point $t_2 > t_1$ in the range of t in which $y(t)>0$. Now look for the solution of the auxiliary equation

$$\frac{d\varphi}{dt} = L\varphi \qquad\qquad \varphi(t_2) = y(t_2)$$

which is solved by

$$\varphi(t) = y(t_2)e^{L(t-t_2)} \tag{2.22}$$

At $t=t_2$ we have

$$\frac{dy}{dt}(t_2) \leq L\,y(t_2) = L\varphi(t_2) = \frac{d\varphi}{dt}(t_2) \tag{2.23}$$

That is, at $t=t_2$, the slope of y is smaller than that of the exponential $\varphi(t)$ and the graphs of the two should look as in Fig. 2.2. However, y vanishes at t_1 while φ(t) only vanishes when $t \to -\infty$. Hence, the two functions must intersect again at some point $t_1 < t_3 < t_2$. At t_3 the slope of φ must be smaller than that of y, in contradiction to Eq. (2.16). Consequently, the assumption that $y=x_1(t)-x_2(t)$ can be nonzero is wrong and the two solutions must be identical. Hence, uniqueness of the solution is a consequence of the Lipschitz condition.

Another way to look at this is to note that it takes the exponential function $\varphi(t)$ an infinite amount of time to reach zero from any nonzero initial value. A function that varies more slowly than the exponential (here $dy/dt<d\varphi/dt$) can certainly not reach zero in a finite amount of time. Hence, the contradiction noted above. Put more formally, the time required for y to reach zero from a

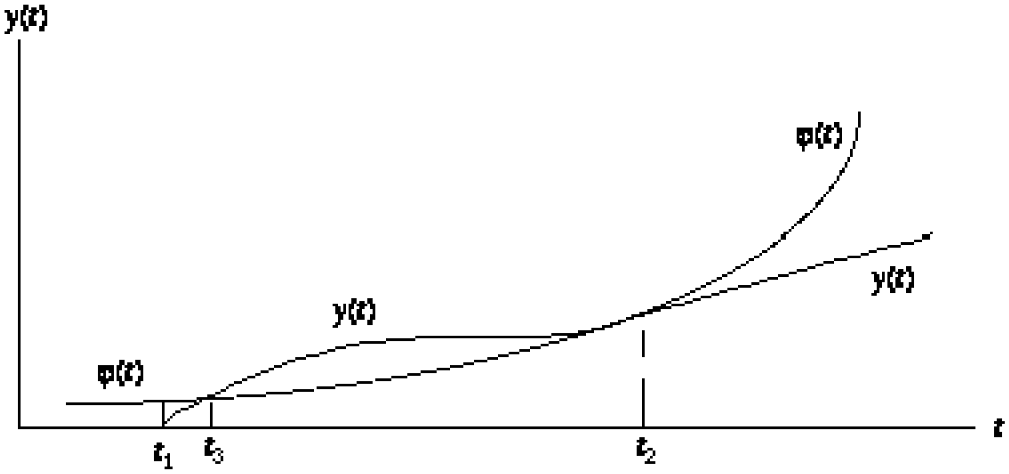


Fig. 2.2

nonzero point y_0 is given by

$$\Delta t = \int_0^{y_0} \frac{dy}{\left[f(x_1(t)) - f(x_2(t))\right]} \geq \int_0^{y_0} \frac{dy}{L\,y} = \infty \tag{2.24}$$

where the Lipschitz condition has been used. If the condition is not satisfied, Δt may be finite. As an example, consider the following case:

$$\frac{dx}{dt} = x^{\alpha} \qquad 0 < \alpha < 1 \tag{2.25}$$

which is solved by

$$x(t) = \begin{cases} 0 & 0 \leq t \leq 1 \\ \left[(1-\alpha)(t-t_0)\right]^{\frac{1}{1-\alpha}} & t > t_0 \end{cases} \tag{2.26}$$

for *any* $t_0 \geq 0$. That is, an infinite number of solutions exists. The reason is that for $|x| < 1$ *there is no constant* L for which

$$\left|x^{\alpha}\right| \leq L|x|$$

The time required to reach x=0 from any finite point is

$$\Delta t = \int_0^{x_0} \frac{ds}{x^{\alpha}} = \frac{1}{1-\alpha}\, x_0^{\;1-\alpha} < \infty \tag{2.27}$$

2.4 The Gronwall lemma [7]

Let $a(t)$, $b(t)$, $c(t) \geq 0$ be continuous in some interval (t_0,t_1) and $Z(t)$ be continuous and satisfying

$$|Z(t)| \leq a(t)\int_{t_0}^{t} b(s)|Z(s)|ds + c(t) \qquad t_1 \leq t_0, \leq t_2 t \tag{2.28}$$

then

$$|Z(t)| \le a(t)\int_{t_0}^{t} b(s)c(s)\exp\left(\int_{s}^{t} a(s')b(s')\,ds'\right)ds + c(t) \tag{2.29}$$

Proof

Define

$$S(t) \equiv \int_{t_0}^{t} b(s)|Z(s)|\,ds \tag{2.30}$$

Then

$$|Z(t)| \le a(t)S(t) + c(t) \tag{2.31}$$

and

$$\frac{dS}{dt} = b(t)\,|Z(t)| \le a(t)b(t)\,S(t) + b(t)c(t) \qquad\qquad S(t_0) = 0 \tag{2.32}$$

$$\frac{dS}{dt} - a\,bS \le bc$$

$$\Rightarrow \frac{d}{dt}\left\{\exp\left(-\int_{t_0}^{t} a(s)b(s)\,ds\right)S(t)\right\} \le \exp\left(-\int_{t_0}^{t} a(s)b(s)\,ds\right)b(t)c(t) \tag{2.33}$$

This leads to

$$\exp\left(-\int_{t_0}^{t} a(s)b(s)\,ds\right)S(t) \le \int_{t_0}^{t} \exp\left(-\int_{t_0}^{s} a(s')b(s')\,ds'\right)b(s)c(s)\,ds$$

$$\Rightarrow S(t) \le \int_{t_0}^{t} \exp\left(-\int_{s}^{t} a(s')b(s')\,ds'\right)b(s)c(s)\,ds \tag{2.34}$$

$$|Z(t)| \le \left[(\delta_2/\delta_1) + \delta_3\right]\exp\left(\delta_1(t - t_0)\right) - (\delta_2/\delta_1)$$

$$|Z(t)| \le a(t)\int_{t_0}^{t} b(s)c(s)\exp\left(\int_{s}^{t} a(s')b(s')\,ds'\right)ds + c(t)$$

A special case of great usefulness is

$$a(t) = \delta_1 \qquad b(t) = 1 \qquad c(t) = \delta_2(t - t_0) + \delta_3$$

$$|Z(t)| \le \left[(\delta_2/\delta_1) + \delta_3\right]\exp\left(\delta_1(t - t_0)\right) - (\delta_2/\delta_1) \tag{2.35}$$

with δ_1, δ_2, and δ_3 constants.

Why has such a simple theorem received a special name (in some textbooks it is given as an exercise to the reader)? The reason is that it has turned out to be extremely useful and powerful. It is the main theorem used in estimating errors incurred in perturbation expansions.

Example: Proof of uniqueness of solution to Volterra integral equations.

Consider the Volterra integral equation

$$x(t) = g(t) + \int_{t_0}^{t} K(t,s)\,x(s)\,ds \tag{2.36}$$

and assume that there are two solutions $x_1(t)$ and $x_2(t)$. The difference $Z(t) \equiv x_1(t) - x_2(t)$ between the two solutions satisfies

$$Z(t) = \int_{t_0}^{t} K(t,s)\,Z(s)\,ds \Rightarrow$$

$$|Z(t)| \le \int_{t_0}^{t} |K(t,s)|\,|Z(s)|\,ds \tag{2.37}$$

If the Kernel $K(t,s)$ is bounded, $|K(t,s)| \le M$, then

$$|Z(t)| \le M \int_{t_0}^{t} |Z(s)|\,ds \tag{2.38}$$

This is a special case of the Gronwall lemma with $a=M$, $b=1$ and $c=0$, yielding

$$|Z(t)| \le 0 \Rightarrow |Z(t)| \equiv 0 \Rightarrow x_1(t) = x_2(t)$$

Hence - uniqueness.

Exercises

2.1 Prove the statement after Eq. (2.7).

2.2 Prove Eq. (2.35) *directly,* that is, without substitution in Eq. (2.34).

2.3 For the equation

$$\frac{dx}{dt} = -\varepsilon x \qquad\qquad x(0) = 1$$

a. find an approximation with an $O(\varepsilon)$ error for $t=O(1)$;
b. find an approximation with an $O(\varepsilon^2)$ error for $t=O(1/\varepsilon^3)$.