Chapter 1

Introduction

1.1 Preliminaries

In this note we concern ourselves with the numerical methods only for the first order ordinary differential (ODE) system in the normal form

$$\frac{dy}{dx} = f(x, y) \tag{1.1}$$

where $x \in \mathbb{R}$, $y \in \mathbb{R}^n$ and $f : \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$. This discussion is adequate because a general m-th order ODE

$$\frac{d^m y}{dx^m} = f(x, y, \frac{dy}{dx}, \ldots, \frac{d^{m-1}y}{dx^{m-1}}) \tag{1.2}$$

with $x, y \in \mathbb{R}$ can be reduced to the system (1.1) by defining

$$\begin{aligned}
y_1 &:= y \\
y_2 &:= \frac{dy_1}{dx} = \frac{dy}{dx} \\
y_3 &:= \frac{dy_2}{dx} = \frac{d^2y}{dx^2} \\
&\vdots \\
y_m &:= \frac{dy_{m-1}}{dx} = \frac{d^{m-1}y}{dx^{m-1}}
\end{aligned} \tag{1.3}$$

and thus

$$\frac{d}{dx}\begin{bmatrix} y_1 \\ \vdots \\ y_{m-1} \\ y_m \end{bmatrix} = \begin{bmatrix} y_2 \\ \vdots \\ y_m \\ f(x, y_1, \ldots, y_{m-1}) \end{bmatrix} \tag{1.4}$$

The system (1.1) may process infinitely many solutions. Each solution is call an *integral curve*. In general we can pick out a particular solution by prescribing certain additional conditions.

Definition 1.1.1 *The system (1.1) together with the initial condition*

$$y(x_0) = y_0 \tag{1.5}$$

is called an initial value problem (IVP).

Definition 1.1.2 *The system (1.1) together with the boundary condition*

$$g(y(a), y(b)) = 0 \tag{1.6}$$

where $a \leq x \leq b$ *and* $g : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *is called a two-point boundary value problem.*

Strictly speaking, an initial value problem is a special case of the two-point boundary value problem, i.e., $g(y(a), y(b)) = y(a) - y_0 = 0$. A special case of (1.6) is also deserved mentioning, i.e., the linear boundary condition

$$g(y(a), y(b)) = B_a y(a) + B_b y(b) = \beta \tag{1.7}$$

where $B_a, B_b \in \mathbb{R}^{n\times n}$ and $\beta \in \mathbb{R}^n$ are specified matrices and vector. It is often assumed that $\text{rank}[B_a, B_b] = n$.

The following example illustrates that even with an additional condition prescribed an integral curve is not necessarily uniquely determined.

Example 1.1.1 The IVP

$$\begin{aligned} \frac{dy}{dx} &= -\sqrt{|1 - y^2|} \\ y(0) &= 1 \end{aligned}$$

has at least two solutions $y(x) \equiv 1$ and $y(x) = \cos x$, and hence has infinitely many solutions. For instance, the function

$$y(x) = \begin{cases} 1 & 0 \le x \le \alpha \\ \cos x - \alpha & \alpha \le x \le \alpha + \pi \\ -1 & \alpha + \pi \le x \end{cases}$$

is a solution for any α.

On the other hand, the following example illustrate that sometimes no integral curve exists to satisfy both the differential equation and the additional condition.

Example 1.1.2 Consider the BVP

$$\begin{aligned} u'' + u &= 0, \; 0 < x < \pi \\ u(0) &= 0 \\ u(b) &= \beta. \end{aligned}$$

The general solution to the differential equation with $u(0) = 0$ is $u(x) = c \sin x$ where c is an arbitrary constant. Thus, the above BVP has no solution if $\beta \neq 0$ and an infinite number of solutions if $\beta = 0$.

Apparently, there are two fundamental questions associated with an ODE problem. The existence question concerns whether there is an integral curve satisfying the ODE as well as the "boundary" condition; the uniqueness question concerns the number of solutions. In addition to being theoretically interesting, consideration of these analytic aspects is also central to numerical computation since they are closely related to the so called conditioning of a given problem. Recall that conditioning concerns perturbations in the solution in terms of perturbations in the data. While we are trying to develop a stable algorithm so that the error propagation is under control, the success of a numerical process depends upon the combination of the well-conditioning of the problem and the stability of the algorithm.

1.2 Existence and Uniqueness

A general criterion that guarantees the existence of a solution to an IVP is the the following Cauchy-Peano Theorem.

Theorem 1.2.1 *Suppose $f(x, y)$ is continuous on an open set $D \subset \mathbb{R} \times \mathbb{R}^n$ containing the initial point (x_0, y_0). Then the IVP has a solution so long as $(x, y(x)) \in D$.*

Note that the Cauchy-Peano Theorem does not have any saying about the uniqueness of a solution as we have seen from Example 1.1.1. Note also that the Cauchy-Peano Theorem concerns only an IVP, not a BVP, as we have seen from Example 1.1.2. A fairly mild condition to guarantee the uniqueness of a solution for an IVP is the following Lipschitz condition.

Definition 1.2.1 *A function $f(x, y)$ is said to satisfy a Lipschitz condition in y with constant L in a region $D \subset \mathbb{R} \times \mathbb{R}^n$ if*

$$\|f(x, u) - f(x, v)\| \leq L\|u - v\| \tag{1.8}$$

for all $(x, u), (x, v) \in D$.

Theorem 1.2.2 *Suppose that $f(x, y)$ and $g(x, y)$ are continuous on an open set $D \in \mathbb{R} \times \mathbb{R}^n$ and that $f(x, y)$ satisfies a Lipschitz condition with constant L. Suppose further that*

$$\|f(x, y) - g(x, y)\| \leq \epsilon$$

for all $(x, y) \in D$. If functions $u(x)$ and $v(x)$ satisfy, respectively, the systems

$$\begin{aligned} \frac{du}{dx} &= f(x, u) \\ \frac{dv}{dx} &= g(x, u) \end{aligned}$$

for $x \in [a, b]$, and $(x, u(x)), (x, v(x)) \in D$, then

$$\|u(x) - v(x)\| \leq \{\|u(a) - v(a)\| + (b - a)\epsilon\} e^{L(x-a)}. \tag{1.9}$$

Proof. By integration, we have

$$u(x) = u(a) + \int_a^x u' dt = u(a) + \int_a^x f(t, u(t))dt$$

and a similar expression for $v(x)$. It follows that

$$\begin{aligned} u(x) - v(x) = u(a) - v(a) \quad &+ \int_a^x \{f(t, v(t)) - g(t, v(t))\} dt \\ &+ \int_a^x \{f(t, u(t)) - f(t, v(t))\} dt. \end{aligned}$$

Upon taking the norm on both sides, we obtain

$$\|u(x) - v(x)\| \leq \|u(a) - v(a)\| + (b - a)\epsilon + \int_a^x L\|u(t) - v(t)\|dt$$

for $a \leq x \leq b$. Define

$$\begin{aligned} \Delta(x) &:= \|u(x) - v(x)\| \\ R(x) &:= \|u(a) - v(a)\| + (b-a)\epsilon + \int_a^x L\|u(t) - v(t)\| dt. \end{aligned}$$

Then

$$\begin{aligned} \Delta(x) &\leq R(x) \\ R' &= L\Delta(x) \leq LR(x) \end{aligned}$$

It follows that

$$\frac{d}{dx}(R(x)e^{-L(x-a)} = e^{-L(x-a)}(R'(x) - LR(x)) \leq 0.$$

The results

$$\begin{aligned} R(x)e^{-L(x-a)} &\leq R(a) \\ \Delta(x) \leq R(a)e^{L(x-a)} & \end{aligned}$$

follows form the non-increasing property. □

Corollary 1.2.3 *Suppose $f(x, y)$ satisfies the condition in Theorem 1.2.2. Then the IVP has a unique solution.*

Remark 1.2.1 Theorem 1.2.1 states that Lipschitz problems are *well-posed*. That is, small perturbations in the stated problem only leads to small changes in the solution. This is the basis of numerical IVP.

In contrast to the IVP, it is in general extremely difficult to provide an explicit existence and uniqueness results for the BVP. The difficulty can be seen from the following argument that forms the basis of one of the numerical methods, called the *shooting method*.

Consider the BVP

$$\begin{aligned} \frac{dy}{dt} &= f(x,y), \; a < x < b & (1.10) \\ g(y(a), y(b)) &= 0. & (1.11) \end{aligned}$$

We associate the related IVP:

$$\begin{aligned} \frac{dw}{dx} &= f(x,w) & (1.12) \\ w(a) &= s & (1.13) \end{aligned}$$

where s is a parameter vector to be described later. Suppose $f(x, w)$ satisfies the Lipschitz condition in w. Then for each initial value $s \in \mathbb{R}^n$ there exists a unique solution for (1.12), denoted by $w(x; s)$ to emphasize the dependence on the parameter s. Define

$$\Phi(s) := g(s, w(b; s)). \tag{1.14}$$

If we could find s such that $\Phi(s) = 0$, then the solution $w(x; s)$ of the IVP would also solve the corresponding BVP, and vise versa. Indeed, we have the following theorem.

Theorem 1.2.4 *Suppose that $f(x, y)$ is continuous on the strip $D = [a, b] \times \mathbb{R}^n$ and satisfies a Lipschitz condition in y. Then the BVP (1.10) and (1.11) has as many solutions as there are distinct roots of $\Phi(s)$.*

Remark 1.2.2 Indeed in order that $\Phi(s)$ be well-defined by (1.14) it must be true that, for the s in question, the solution of the IVP (1.12) and (1.13) exist on the entire interval $[a, b]$. This and some smoothness of $w(x; s)$ are consequences of standard results from ODE theory (see Coddington and Levinson [Chapter1, §7]).

The equation $\Phi(s) = 0$ in general is a system of n nonlinear algebraic equations for the n unknowns s. It is expected that in general such a system may have none, one or many solutions. We learn from the above theorem that perhaps what is important in the BVP is not the over all uniqueness, but rather the *separation* of one solution from the others. More precisely, we define the following concept.

Definition 1.2.2 *A solution $y(x)$ of a BVP is said to be isolated (or locally unique) if there exists a $\rho > 0$ such that in the class of functions*

$$\{z(x) \in C[a, b] \mid \|z - y\|_\infty \leq \rho\},$$

where $\|z - y\|_\infty := \sup_{x \in [a,b]} |z(x) - y(x)|$, $y(x)$ is the only solution.

The isolation of a solution to a general nonlinear BVP can be identified through the corresponding variation equation that in some sense measures the sensitivity of the solution $y(x)$ to a parameter s. We briefly introduce the ideas below: Suppose the solution $y(x)$ depends upon some parameter s so that we denote $y(x) = y(x; s)$. Define

$$z(t) := \frac{\partial y}{\partial s}. \tag{1.15}$$

Then $z(t)$ solves, under some smoothness conditions, the linear BVP:

$$\frac{dz}{dt} = \frac{\partial f(x, y(x))}{\partial y} z \tag{1.16}$$

$$\frac{\partial g(y(a), y(b))}{\partial y(a)} z(a) + \frac{\partial g(y(a), y(b))}{\partial y(b)} z(b) = 0. \tag{1.17}$$

It can be shown that

Theorem 1.2.5 *Suppose that the variational problem (1.16) and (1.17) has a unique solution $z(x) \equiv 0$. Then the solution to the BVP (1.10) and (1.11) is isolated.*

Remark 1.2.3 Suppose the parameter s is taken to represent the initial value as in (1.13). Then an isolated solution of the BVP implies that the corresponding s is a simple root of $\Phi(s) = 0$. Thus, in principle, an isolated solution can be located numerically by solving the equation $\Phi(s) = 0$ by, say, the Newton method.

It should be kept in mind that the linear BVP

$$\frac{dy}{dx} = A(x)y + q(x) \tag{1.18}$$

$$B_a y(a) + B_b y(b) = \beta, \tag{1.19}$$

always serves as the basis of many more complicated theory. For instance, the existence and uniqueness questions for the above linear model can be answered easily. To see this, we apply the superposition principle as follows: Let $y_p(x)$ denote a particular solution to the IVP

$$\frac{dy_p}{dt} = A(x)y_p + q(x) \tag{1.20}$$

$$y_p(a) = y_0. \tag{1.21}$$

For $\nu = 1, \ldots, n$, let $y_\nu(x)$ denote a solution of the homogeneous IVP

$$\frac{dy_\nu}{dt} = A(x)y_\nu \tag{1.22}$$

$$y_\nu(a) = e_\nu \tag{1.23}$$

where e_ν is the ν-th standard unit vector in $\mathbb{R}^n$. Then a general solution $y(x)$ to the differential equation (1.18) is of the form

$$y(x) = y_p(x) + \sum_{\nu=1}^{n} \xi_\nu y_\nu(x) \tag{1.24}$$

where $\xi_\nu, \nu = 1, \ldots, n$ are arbitrary. Introducing the fundamental solution matrix

$$Y(x) := [y_1(x), \ldots, y_n(x)], \tag{1.25}$$

then to satisfy the boundary condition we must solve the equation

$$[B_a + B_b Y(b)]\xi = \beta - B_a y_0 - B_b y_p(b) \tag{1.26}$$

for $\xi := [\xi_1, \ldots, \xi_n]^T$. Denote the coefficient matrix

$$Q := B_a + B_b Y(b). \tag{1.27}$$

We have the following theorem.

Theorem 1.2.6 *Let $A(x), q(x) \in C^{p+1}[a,b]$. Then the BVP (1.18) and (1.19) has a unique solution $y(x) \in C^{p+1}[a,b]$ if and only if Q is nonsingular.*