MATH 537, Fall 2020
Ordinary Differential Equations

Lecture #24-Supp

Fixed Point Iteration

Instructor: Dr. Bo-Wen Shen*
Department of Mathematics and Statistics
San Diego State University

Computational Ordinary Differential Equations

Lecture Notes #19_0 --- Fixed Point Iteration

Instructor: Dr. Bo-Wen Shen

sdsu.math542.shen@gmail.com

Department of Mathematics and Statistics

San Diego State University

http://bwshen.sdsu.edu

Fall 2019

References

Textbooks

1. Burden R., D. Faires, and A. Burden, 2014 : Numerical Analysis. 10th edition. Cengage.
2. Stewart, Multivariable Calculus.

Fixed Point and Fixed Point Theorem

A *fixed point* for a function is a number at which the value of the function does not change when the function is applied.

The number p is a **fixed point** for a given function g if $g(p) = p$.

In this section we consider the problem of finding solutions to fixed-point problems and the connection between the fixed-point problems and the root-finding problems we wish to solve. Root-finding problems and fixed-point problems are equivalent classes in the following sense:

- Given a root-finding problem $f(p) = 0$, we can define functions g with a fixed point at p in a number of ways, for example, as

$$g(x) = x - f(x) \quad \text{or as} \quad g(x) = x + \boxed{3f(x).}$$

- Conversely, if the function g has a fixed point at p, then the function defined by

$$f(x) = x - g(x)$$

has a zero at p.

Burden et al. (2014)

A Note:

- Given $f(x)$, we can define $\mathrm{g(x)} = f(x) + x$.
- Thus, a fixed point x_o of g(x), $\mathrm{i.e.,}$ $g(x_o) = x_o$ is a zero of $f(x)$, $\mathrm{i.e.,}$ $f(x_o) = 0$.

➢ Given $f(x)$, we can define $\mathrm{g(x)}$ as follows:

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

➢ Thus, a fixed point x_o of g(x), $\mathrm{i.e.,}$ $g(x_o) = x_o$ is a zero of $f(x)$, $\mathrm{i.e.,}$ $f(x_o) = 0$.

Fixed Point and Fixed Point Theorem

To approximate the fixed point of a function g, we choose an initial approximation p_0 and generate the sequence $\{p_n\}_{n=0}^{\infty}$ by letting $p_n = g(p_{n-1})$, for each $n \geq 1$. If the sequence converges to p and g is continuous, then

$$p = \lim_{n\to\infty} p_n = \lim_{n\to\infty} g(p_{n-1}) = g\left(\lim_{n\to\infty} p_{n-1}\right) = g(p),$$

and a solution to $x = g(x)$ is obtained. This technique is called **fixed-point**, or **functional iteration**. The procedure is illustrated in Figure 2.7 and detailed in Algorithm 2.2.

Fixed Point Theorem

Theorem 2.4 (Fixed-Point Theorem)

Let $g \in C[a,b]$ be such that $g(x) \in [a,b]$, for all x in $[a,b]$. Suppose, in addition, that g' exists on (a,b) and that a constant $0 < k < 1$ exists with

$$\boxed{|g'(x)| \le k,} \quad \text{for all } x \in (a,b).$$

Then for any number p_0 in $[a,b]$, the sequence defined by

$$p_n = g(p_{n-1}), \quad n \ge 1,$$

converges to the unique fixed point p in $[a,b]$. ■

condition for convergence: $|g'(x) < 1|$

Fixed Point Theorem: A Proof

Proof Theorem 2.3 implies that a unique point p exists in $[a, b]$ with $g(p) = p$. Since g maps $[a, b]$ into itself, the sequence $\{p_n\}_{n=0}^{\infty}$ is defined for all $n \geq 0$, and $p_n \in [a, b]$ for all n. Using the fact that $|g'(x)| \leq k$ and the Mean Value Theorem 1.8, we have, for each n,

$$|p_n - p| = |g(p_{n-1}) - g(p)| = |g'(\xi_n)||p_{n-1} - p| \leq k|p_{n-1} - p|,$$

where $\xi_n \in (a, b)$. Applying this inequality inductively gives

$$|p_n - p| \leq k|p_{n-1} - p| \leq k^2|p_{n-2} - p| \leq \cdots \leq k^n|p_0 - p|. \tag{2.4}$$

Since $0 < k < 1$, we have $\lim_{n\to\infty} k^n = 0$ and

$$\lim_{n\to\infty} |p_n - p| \leq \lim_{n\to\infty} k^n|p_0 - p| = 0.$$

Hence $\{p_n\}_{n=0}^{\infty}$ converges to p. ■ ■ ■

Newton's Method

Newton's method $$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, \quad \text{for } n \geq 1. \tag{2.7}$$

Newton's method is a functional iteration technique with $p_n = g(p_{n-1})$, for which

$$g(p_{n-1}) = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}, \quad \text{for } n \geq 1. \tag{2.11}$$

In fact, this is the functional iteration technique that was used to give the rapid convergence we saw in column (**e**) of Table 2.2 in Section 2.2.

It is clear from Equation (2.7) that Newton's method cannot be continued if $f'(p_{n-1}) = 0$ for some n. In fact, we will see that the method is most effective when f' is bounded away from zero near p.

When g(p*)=p*, eq. 2.11 leads to f(p*)=0. p* is a solution of f(x).

Convergence using Newton's Method

$$\boxed{|g'(x)| \leq k,} \quad \text{for all } x \in (a, b).$$

Theorem 2.6 Let $f \in C^2[a, b]$. If $p \in (a, b)$ is such that $f(p) = 0$ and $f'(p) \neq 0$, then there exists a $\delta > 0$ such that Newton's method generates a sequence $\{p_n\}_{n=1}^{\infty}$ converging to p for any initial approximation $p_0 \in [p - \delta, p + \delta]$. ■

1. Theorem 2.6 states that, under reasonable assumptions, Newton's method converges provided a sufficiently accurate initial approximation is chosen.
2. It also implies that the constant k that bounds the derivative of g , and, consequently, indicates the speed of convergence of the method, decreases to 0 as the procedure continues.
3. This result is important for the theory of Newton's method, but it is seldom applied in practice because it does not tell us how to determine δ.

Convergence using Newton's Method: A Proof

Proof The proof is based on analyzing Newton's method as the functional iteration scheme $p_n = g(p_{n-1})$, for $n \geq 1$, with

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

Let k be in $(0, 1)$. We first find an interval $[p-\delta, p+\delta]$ that g maps into itself and for which $\boxed{|g'(x)| \leq k}$, for all $x \in (p-\delta, p+\delta)$.

Since f' is continuous and $f'(p) \neq 0$, part (**a**) of Exercise 29 in Section 1.1 implies that there exists a $\delta_1 > 0$, such that $f'(x) \neq 0$ for $x \in [p-\delta_1, p+\delta_1] \subseteq [a, b]$. Thus g is defined and continuous on $[p-\delta_1, p+\delta_1]$. Also

$$\boxed{g'(x) = 1 - \frac{f'(x)f'(x) - f(x)f''(x)}{[f'(x)]^2} = \frac{f(x)f''(x)}{[f'(x)]^2},}$$

for $x \in [p-\delta_1, p+\delta_1]$, and, since $f \in C^2[a, b]$, we have $g \in C^1[p-\delta_1, p+\delta_1]$.

By assumption, $f(p) = 0$, so

$$\boxed{g'(p) = \frac{f(p)f''(p)}{[f'(p)]^2} = 0.}$$

Since g' is continuous and $0 < k < 1$, part (**b**) of Exercise 29 in Section 1.1 implies that there exists a δ, with $0 < \delta < \delta_1$, and

$$|g'(x)| \leq k, \quad \text{for all } x \in [p-\delta, p+\delta].$$

A System of Nonlinear Equations

A system of nonlinear equations has the form

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= 0, \\ f_2(x_1, x_2, \ldots, x_n) &= 0, \\ \vdots \qquad\qquad & \quad \vdots \\ f_n(x_1, x_2, \ldots, x_n) &= 0, \end{aligned} \tag{10.1}$$

where each function f_i can be thought of as mapping a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ of the n-dimensional space $\mathbb{R}^n$ into the real line $\mathbb{R}$. A geometric representation of a nonlinear system when $n = 2$ is given in Figure 10.1.

This system of n nonlinear equations in n unknowns can also be represented by defining a function $\mathbf{F}$ mapping $\boxed{\mathbb{R}^n \text{ into } \mathbb{R}^n}$ as

$$\mathbf{F}(x_1, x_2, \ldots, x_n) = (\boxed{f_1}(x_1, x_2, \ldots, x_n), \boxed{f_2}(x_1, x_2, \ldots, x_n), \ldots, \boxed{f_n}(x_1, x_2, \ldots, x_n))^t.$$

If vector notation is used to represent the variables $x_1, x_2, \ldots, x_n$, then system (10.1) assumes the form

$$\mathbf{F}(\mathbf{x}) = \mathbf{0}. \tag{10.2}$$

The functions $f_1, f_2, \ldots, f_n$ are called the **coordinate functions** of $\mathbf{F}$.

High-dimensional Fixed Point

In Chapter 2, an iterative process for solving an equation $f(x) = 0$ was developed by first transforming the equation into the fixed-point form $x = g(x)$. A similar procedure will be investigated for functions from $\mathbb{R}^n$ into $\mathbb{R}^n$.

Definition 10.5 A function $\mathbf{G}$ from $D \subset \mathbb{R}^n$ into $\mathbb{R}^n$ has a **fixed point** at $\mathbf{p} \in D$ if $\boxed{\mathbf{G}(\mathbf{p}) = \mathbf{p}}$. ■

The following theorem extends the Fixed-Point Theorem 2.4 on page 62 to the n-dimensional case. This theorem is a special case of the Contraction Mapping Theorem, and its proof can be found in [Or2], p. 153.

Theorem 10.6 Let $D = \{(x_1, x_2, \ldots, x_n)^t \mid a_i \le x_i \le b_i$, for each $i = 1, 2, \ldots, n\}$ for some collection of constants $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$. Suppose $\mathbf{G}$ is a continuous function from $D \subset \mathbb{R}^n$ into $\mathbb{R}^n$ with the property that $\mathbf{G}(\mathbf{x}) \in D$ whenever $\mathbf{x} \in D$. Then $\mathbf{G}$ has a fixed point in D.

Moreover, suppose that all the component functions of $\mathbf{G}$ have continuous partial derivatives and a constant $K < 1$ exists with

$$\left| \frac{\partial g_i(\mathbf{x})}{\partial x_j} \right| \le \frac{K}{n}, \quad \text{whenever } \mathbf{x} \in D,$$

for each $j = 1, 2, \ldots, n$ and each component function g_i. Then the sequence $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ defined by an arbitrarily selected $\mathbf{x}^{(0)}$ in D and generated by

$$\mathbf{x}^{(k)} = G(\mathbf{x}^{(k-1)}), \quad \text{for each } k \ge 1$$

converges to the unique fixed point $\mathbf{p} \in D$ and

$$\left\| \mathbf{x}^{(k)} - \mathbf{p} \right\|_\infty \le \frac{K^k}{1-K} \left\| \mathbf{x}^{(1)} - \mathbf{x}^{(0)} \right\|_\infty . \tag{10.3}$$

■

Newton's Method in the N-D Space

1-D

$$g(x) = x - \frac{f(x)}{f'(x)}$$

A similar approach in the n-dimensional case involves a matrix

N-D

$$A(\mathbf{x}) = \begin{bmatrix} a_{11}(\mathbf{x}) & a_{12}(\mathbf{x}) & \cdots & a_{1n}(\mathbf{x}) \\ a_{21}(\mathbf{x}) & a_{22}(\mathbf{x}) & \cdots & a_{2n}(\mathbf{x}) \\ \vdots & \vdots & & \vdots \\ a_{n1}(\mathbf{x}) & a_{n2}(\mathbf{x}) & \cdots & a_{nn}(\mathbf{x}) \end{bmatrix}, \tag{10.5}$$

where $\boxed{\text{each}}$ of the entries $a_{ij}(\mathbf{x})$ is a function from $\boxed{\mathbb{R}^n \text{ into } \mathbb{R}}$. This requires that $A(\mathbf{x})$ be found so that

$$\mathbf{G}(\mathbf{x}) = \mathbf{x} - A(\mathbf{x})^{-1}\mathbf{F}(\mathbf{x})$$

gives quadratic convergence to the solution of $\mathbf{F}(\mathbf{x}) = \mathbf{0}$, assuming that $\boxed{A(\mathbf{x}) \text{ is nonsingular}}$ at the fixed point $\mathbf{p}$ of $\mathbf{G}$.

The following theorem parallels Theorem 2.8 on page 80. Its proof requires being able to express $\mathbf{G}$ in terms of its Taylor series in n variables about $\mathbf{p}$.

Newton's Method in the N-D Space

$$1D:\ g(x) = x - \frac{f(x)}{f'(x)}$$

The Jacobian Matrix

Define the matrix $J(\mathbf{x})$ by

$$J(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}) & \frac{\partial f_n}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial f_n}{\partial x_n}(\mathbf{x}) \end{bmatrix}. \tag{10.8}$$

Then conditions (10.6) and (10.7) require that

$$A(\mathbf{p})^{-1}J(\mathbf{p}) = I, \text{the identity matrix,} \quad \text{so } A(\mathbf{p}) = J(\mathbf{p}).$$

An appropriate choice for $A(\mathbf{x})$ is, consequently, $A(\mathbf{x}) = J(\mathbf{x})$ since this satisfies condition (iii) in Theorem 10.7. The function $\mathbf{G}$ is defined by

$$\mathbf{G}(\mathbf{x}) = \mathbf{x} - J(\mathbf{x})^{-1}\mathbf{F}(\mathbf{x}),$$

and the functional iteration procedure evolves from selecting $\mathbf{x}^{(0)}$ and generating, for $k \geq 1$,

$$\mathbf{x}^{(k)} = \mathbf{G}\left(\mathbf{x}^{(k-1)}\right) = \mathbf{x}^{(k-1)} - J\left(\mathbf{x}^{(k-1)}\right)^{-1}\mathbf{F}\left(\mathbf{x}^{(k-1)}\right). \tag{10.9}$$

This is called **Newton's method for nonlinear systems**, and it is generally expected