Jim Lambers
MAT 460
Fall Semester 2009-10
Lecture 3 Notes

These notes correspond to Section 1.1 in the text.

Review of Calculus, cont'd

Riemann Sums and the Definite Integral

There are many cases in which some quantity is defined to be the product of two other quantities. For example, a rectangle of width w has uniform height h, and the area A of the rectangle is given by the formula $A = wh$. Unfortunately, in many applications, we cannot necessarily assume that certain quantities such as height are constant, and therefore formulas such as $A = wh$ cannot be used directly. However, they can be used indirectly to solve more general problems by employing the notation known as *integral calculus*.

Suppose we wish to compute the area of a shape that is not a rectangle. To simplify the discussion, we assume that the shape is bounded by the vertical lines $x = a$ and $x = b$, the x-axis, and the curve defined by some continuous function $y = f(x)$, where $f(x) \geq 0$ for $a \leq x \leq b$. Then, we can approximate this shape by n rectangles that have width $\Delta x = (b - a)/n$ and height $f(x_i)$, where $x_i = a + i\Delta x$, for $i = 0, \ldots, n$. We obtain the approximation

$$A \approx A_n = \sum_{i=1}^{n} f(x_i)\Delta x.$$

Intuitively, we can conclude that as $n \to \infty$, the approximate area A_n will converge to the exact area of the given region. This can be seen by observing that as n increases, the n rectangles defined above comprise a more accurate approximation of the region.

More generally, suppose that for each $n = 1, 2, \ldots$, we define the quantity R_n by choosing points $a = x_0 < x_1 < \cdots < x_n = b$, and computing the sum

$$R_n = \sum_{i=1}^{n} f(x_i^*)\Delta x_i, \quad \Delta x_i = x_i - x_{i-1}, \quad x_{i-1} \leq x_i^* \leq x_i.$$

The sum that defines R_n is known as a *Riemann sum*. Note that the interval $[a, b]$ need not be divided into subintervals of equal width, and that $f(x)$ can be evaluated at *arbitrary* points belonging to each subinterval.

If $f(x) \geq 0$ on $[a, b]$, then R_n converges to the area under the curve $y = f(x)$ as $n \to \infty$, provided that the widths of all of the subintervals $[x_{i-1}, x_i]$, for $i = 1, \ldots, n$, approach zero. This behavior is ensured if we require that

$$\lim_{n\to\infty} \delta(n) = 0, \quad \text{where} \quad \delta(n) = \max_{1\le i\le n} \Delta x_i.$$

This condition is necessary because if it does not hold, then, as $n \to \infty$, the region formed by the n rectangles will not converge to the region whose area we wish to compute. If f assumes negative values on $[a, b]$, then, under the same conditions on the widths of the subintervals, R_n converges to the *net* area between the graph of f and the x-axis, where area below the x-axis is counted negatively.

We define the *definite integral* of $f(x)$ from a to b by

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} R_n,$$

where the sequence of Riemann sums $\{R_n\}_{n=1}^{\infty}$ is defined so that $\delta(n) \to 0$ as $n \to \infty$, as in the previous discussion. The function $f(x)$ is called the *integrand*, and the values a and b are the *lower* and *upper limits of integration*, respectively. The process of computing an integral is called *integration*.

In this course, we will study the problem of computing an approximation to the definite integral of a given function $f(x)$ over an interval $[a, b]$. We will learn a number of techniques for computing such an approximation, and all of these techniques involve the computation of an appropriate Riemann sum.

Extreme Values

In many applications, it is necessary to determine where a given function attains its minimum or maximum value. For example, a business wishes to maximize profit, so it can construct a function that relates its profit to variables such as payroll or maintenance costs. We now consider the basic problem of finding a maximum or minimum value of a general function $f(x)$ that depends on a single independent variable x. First, we must precisely define what it means for a function to *have* a maximum or minimum value.

Definition *(Absolute extrema) A function f has a* **absolute maximum** *or* **global maximum** *at c if $f(c) \ge f(x)$ for all x in the domain of f. The number $f(c)$ is called the* **maximum value** *of f on its domain. Similarly, f has a* **absolute minimum** *or* **global minimum** *at c if $f(c) \le f(x)$ for all x in the domain of f. The number $f(c)$ is then called the* **minimum value** *of f on its domain. The maximum and minimum values of f are called the* **extreme values** *of f, and the absolute maximum and minimum are each called an* **extremum** *of f.*

Before computing the maximum or minimum value of a function, it is natural to ask whether it is possible to determine in advance whether a function even has a maximum or minimum, so that effort is not wasted in trying to solve a problem that has no solution. The following result is very helpful in answering this question.

Theorem *(Extreme Value Theorem) If f is continuous on $[a, b]$, then f has an absolute maximum and an absolute minimum on $[a, b]$.*

Now that we can easily determine whether a function has a maximum or minimum on a closed interval $[a, b]$, we can develop an method for actually finding them. It turns out that it is easier to find points at which f attains a maximum or minimum value in a "local" sense, rather than a "global" sense. In other words, we can best find the absolute maximum or minimum of f by finding points at which f achieves a maximum or minimum with respect to "nearby" points, and then determine which of these points is the absolute maximum or minimum. The following definition makes this notion precise.

Definition *(Local extrema) A function f has a* **local maximum** *at c if $f(c) \geq f(x)$ for all x in an open interval containing c. Similarly, f has a* **local minimum** *at c if $f(c) \leq f(x)$ for all x in an open interval containing c. A local maximum or local minimum is also called a* **local extremum***.*

At each point at which f has a local maximum, the function either has a horizontal tangent line, or no tangent line due to not being differentiable. It turns out that this is true in general, and a similar statement applies to local minima. To state the formal result, we first introduce the following definition, which will also be useful when describing a method for finding local extrema.

Definition*(Critical Number) A number c in the domain of a function f is a* **critical number** *of f if $f'(c) = 0$ or $f'(c)$ does not exist.*

The following result describes the relationship between critical numbers and local extrema.

Theorem *(Fermat's Theorem) If f has a local minimum or local maximum at c, then c is a critical number of f; that is, either $f'(c) = 0$ or $f'(c)$ does not exist.*

This theorem suggests that the maximum or minimum value of a function $f(x)$ can be found by solving the equation $f'(x) = 0$. As mentioned previously, we will be learning techniques for solving such equations in this course. These techniques play an essential role in the solution of problems in which one must compute the maximum or minimum value of a function, subject to constraints on its variables. Such problems are called *optimization problems*. Although we will not discuss optimization problems in this course, we will learn about some of the building blocks of methods for solving these very important problems.

The Mean Value Theorem

While the derivative describes the behavior of a function at a point, we often need to understand how the derivative influences a function's behavior on an interval. This understanding is essential in numerical analysis because, it is often necessary to approximate a function $f(x)$ by a function

$g(x)$ using knowledge of $f(x)$ and its derivatives at various points. It is therefore natural to ask how well $g(x)$ approximates $f(x)$ away from these points.

The following result, a consequence of Fermat's Theorem, gives limited insight into the relationship between the behavior of a function on an interval and the value of its derivative at a point.

Theorem *(Rolle's Theorem) If f is continuous on a closed interval $[a, b]$ and is differentiable on the open interval (a, b), and if $f(a) = f(b)$, then $f'(c) = 0$ for some number c in (a, b).*

By applying Rolle's Theorem to a function f, then to its derivative f', its second derivative f'', and so on, we obtain the following more general result, which will be useful in analyzing the accuracy of methods for approximating functions by polynomials.

Theorem *(Generalized Rolle's Theorem) Let $x_0, x_1, x_2, \ldots, x_n$ be distinct points in an interval $[a, b]$. If f is n times differentiable on (a, b), and if $f(x_i) = 0$ for $i = 0, 1, 2, \ldots, n$, then $f^{(n)}(c) = 0$ for some number c in (a, b).*

A more fundamental consequence of Rolle's Theorem is the Mean Value Theorem itself, which we now state.

Theorem *(Mean Value Theorem) If f is continuous on a closed interval $[a, b]$ and is differentiable on the open interval (a, b), then*

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

for some number c in (a, b).

Remark The expression

$$\frac{f(b) - f(a)}{b - a}$$

is the slope of the secant line passing through the points $(a, f(a))$ and $(b, f(b))$. The Mean Value Theorem therefore states that under the given assumptions, the slope of this secant line is equal to the slope of the tangent line of f at the point $(c, f(c))$, where $c \in (a, b)$. $\Box$

The Mean Value Theorem has the following practical interpretation: the average rate of change of $y = f(x)$ with respect to x on an interval $[a, b]$ is equal to the instantaneous rate of change y with respect to x at some point in (a, b).

The Mean Value Theorem for Integrals

Suppose that $f(x)$ is a continuous function on an interval $[a, b]$. Then, by the Fundamental Theorem of Calculus, $f(x)$ has an antiderivative $F(x)$ defined on $[a, b]$ such that $F'(x) = f(x)$. If we apply the Mean Value Theorem to $F(x)$, we obtain the following relationship between the integral of f over $[a, b]$ and the value of f at a point in (a, b).

Theorem *(Mean Value Theorem for Integrals) If f is continuous on $[a, b]$, then*

$$\int_a^b f(x)\,dx = f(c)(b-a)$$

for some c in (a, b).

In other words, f assumes its average value over $[a, b]$, defined by

$$f_{ave} = \frac{1}{b-a}\int_a^b f(x)\,dx,$$

at some point in $[a, b]$, just as the Mean Value Theorem states that the derivative of a function assumes its average value over an interval at some point in the interval.

The Mean Value Theorem for Integrals is also a special case of the following more general result.

Theorem *(Weighted Mean Value Theorem for Integrals) If f is continuous on $[a, b]$, and g is a function that is integrable on $[a, b]$ and does not change sign on $[a, b]$, then*

$$\int_a^b f(x)g(x)\,dx = f(c)\int_a^b g(x)\,dx$$

for some c in (a, b).

In the case where $g(x)$ is a function that is easy to antidifferentiate and $f(x)$ is not, this theorem can be used to obtain an estimate of the integral of $f(x)g(x)$ over an interval.

Example Let $f(x)$ be continuous on the interval $[a, b]$. Then, for any $x \in [a, b]$, by the Weighted Mean Value Theorem for Integrals, we have

$$\int_a^x f(s)(s-a)\,ds = f(c)\int_a^x (s-a)\,ds = f(c)\left.\frac{(s-a)^2}{2}\right|_a^x = f(c)\frac{(x-a)^2}{2},$$

where $a < c < x$. It is important to note that we can apply the Weighted Mean Value Theorem because the function $g(x) = (x-a)$ does not change sign on $[a, b]$. □

Taylor's Theorem

In many cases, it is useful to approximate a given function $f(x)$ by a polynomial, because one can work much more easily with polynomials than with other types of functions. As such, it is necessary to have some insight into the accuracy of such an approximation. The following theorem, which is a consequence of the Weighted Mean Value Theorem for Integrals, provides this insight.

Theorem *(Taylor's Theorem) Let f be n times continuously differentiable on an interval $[a, b]$, and suppose that $f^{(n+1)}$ exists on $[a, b]$. Let $x_0 \in [a, b]$. Then, for any point $x \in [a, b]$,*

$$f(x) = P_n(x) + R_n(x),$$

where

$$\begin{aligned} P_n(x) &= \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!}(x-x_0)^j \\ &= f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2}f''(x_0)(x-x_0)^2 + \cdots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n \end{aligned}$$

and

$$R_n(x) = \int_{x_0}^{x} \frac{f^{(n+1)}(s)}{n!}(x-s)^n\,ds = \frac{f^{(n+1)}(\xi(x))}{(n+1)!}(x-x_0)^{n+1},$$

where $\xi(x)$ *is between* x_0 *and* x.

The polynomial $P_n(x)$ is the nth *Taylor polynomial* of f with *center* x_0, and the expression $R_n(x)$ is called the *Taylor remainder* of $P_n(x)$. When the center x_0 is zero, the nth Taylor polynomial is also known as the nth *Maclaurin polynomial*.

The final form of the remainder is obtained by applying the Mean Value Theorem for Integrals to the integral form. As $P_n(x)$ can be used to approximate $f(x)$, the remainder $R_n(x)$ is also referred to as the *truncation error* of $P_n(x)$. The accuracy of the approximation on an interval can be analyzed by using techniques for finding the extreme values of functions to bound the $(n+1)$-st derivative on the interval.

Because approximation of functions by polynomials is employed in the development and analysis of many techniques in numerical analysis, the usefulness of Taylor's Theorem cannot be overstated. In fact, it can be said that Taylor's Theorem is the Fundamental Theorem of Numerical Analysis, just as the theorem describing inverse relationship between derivatives and integrals is called the Fundamental Theorem of Calculus.