Error estimates and convex and concave functions

Recall the statement of Rolle's Theorem:

Theorem 1 (Rolle's Theorem) *Suppose that f is a continuous function with domain $[a, b]$, that $f(a) = f(b)$ and f is differentiable at every $x \in (a, b)$. Then there is at least one $c \in (a, b)$ where $f'(c) = 0$.*

We will use Rolle's Theorem in a different way to determine the way the curves bend.

1 The Second Rolle's Theorem

We have the following extension of Rolle's theorem to second derivatives.

Corollary 1 (Second Rolle's Theorem) *Suppose that f is continuous on $[a, b]$, $f(b) = f(a) = 0$, and that $f''(x)$ exists for each $x \in (a, b)$. Suppose one of the following is true:*

Case I: *There is some $c \in (a, b)$ such that $f(c) = 0$*

Case II: *$f'(a) = 0$ and f' is continuous at a.*

Case III: *$f'(b) = 0$ and f' is continuous at b.*

Then there is some $d \in (a, b)$ so that $f''(d) = 0$.

Reason: In **Case I** apply Rolle's Theorem to f on $[a, c]$ and $[c, b]$ separately to infer that there are u and v with $a < u < c < v < b$ and $f'(u) = f'(v) = 0$. Observe that Rolle's Theorem now applies to f' on $[u, v]$.

In **Case II** apply Rolle's Theorem on $[a, b]$ to infer that there is some $u \in (a, b)$ so that $f'(u) = 0$. Observe that since $f'(a) = 0$ and f' is continuous on $[a, u]$ that Rolle's Theorem applies to f' on $[a, u]$.

Case III is similar to **Case II** **QED**

1.1 Error Estimate in the Best Linear Approximation

An important application of the Second Rolle's Theorem is error estimation in linear approximations. For example, recall that the best linear approximation to $f(x)$ at $x = a$ is $L(x) = f(a) + f'(a)(x - a)$. For a given number b, how close is $L(b)$ to $f(b)$? Here is one answer.

Theorem 2 *Suppose that f is continuous on $[a, b]$, $f''(x)$ exists for each $x \in (a, b)$, and f' is continuous at a. Let $L(x) = f(a) + f'(a)(x - a)$ be the best linear approximation to $f(x)$ at $x = a$. Then there is some $c \in (a, b)$ so that*

$$f(b) - L(b) = \frac{1}{2} f''(c)(b - a)^2.$$

Reason: Define

$$F(x) = (f(x) - L(x))(b - a)^2 - (f(b) - L(b))(x - a)^2$$

It is easy to check that

$$\begin{aligned}
F(a) &= 0 \\
F(b) &= 0 \\
F'(x) &= (f'(x) - f'(a))(b - a)^2 - 2(f(b) - L(b)(x - a) \\
F'(a) &= 0 \\
F''(x) &= f''(x)(b - a)^2 - 2(f(b) - L(b))
\end{aligned}$$

so it follows from the Second Rolle's Theorem that there is some $c \in (a, b)$ so that

$$\begin{aligned}
0 &= F''(c) \\
&= f''(c)(b - a)^2 - 2(f(b) - L(b)) \\
f(b) - L(b) &= \frac{1}{2} f''(c)(b - a)^2
\end{aligned}$$

as desired. **QED**

Note that the same conclusion holds if $b < a$.

For example, suppose $f(x) = \ln(1+x)$. We know that the best linear approximation to $\ln(1+x)$ at $x = 0$ is $L(x) = x$. We know that $f''(x) = -1/(1+x)^2$ for $x > -1$. Then for some c between 0 and x

$$\ln(1+x) = x - \frac{1}{2}\frac{1}{(1+c)^2}x^2.$$

For example, suppose we want an estimate of $\ln(1.1)$. For $0 < c < 0.1$ we have

$$\frac{1}{1.1^2} \leq \frac{1}{(1+y)^2} \leq 1$$

so

$$0.004 < \frac{1}{242} \leq 0.1 - \ln(1.1) \leq \frac{1}{200} = 0.005$$

or,

$$0.095 < \ln(1.1) < 0.096.$$

Note that to 10 digits $\ln(1.1) \approx 0.9531017980$.

Corollary 2 *If $f''(x) \geq 0$ on an open interval I then for each $a \in I$,*

$$f(x) \geq f(a) + f'(a)(x-a).$$

In other words, the graph of $f(x)$ lies on or above the graph of any tangent line so long as $x \in I$.

Reason: Let $L(x) = f(a) + f'(a)(x-a)$, and the graph of $y = L(x)$ is the graph of the tangent line to $y = f(x)$ at the point $(a, f(a))$. We know from Theorem 2 that

$$f(x) = L(x) + \frac{1}{2}f''(c_x)(x-a)^2 \geq L(x)$$

where c_x lies strictly between a and x. **QED**

Of course the reverse is true if $f''(x) < 0$ on I, that is, the graph lies under the tangent lines.

1.2 Newton's Method

Suppose that we are to solve the equation $f(x) = 0$, and there is no algebraic solution. We know that if f is continuous then we know that we can use the bisection method. If f is differentiable and its second derivative does not change sign, we can do better. The idea is to find some a so that $f(a) \approx 0$. Then instead of trying to solve $f(x) = 0$ we solve $f(a) + f'(a)(x-a) = 0$, that is we look to see where the tangent line at $(a, f(a))$ crosses the horizontal axis. We then repeat this procedure.

Theorem 3 (Newton's Method) *Suppose $a < c < b$, that $f(a) = 0$, and that f is twice differentiable on (a, b), and f and f' are each continuous at a. Assume $f'(x) > 0$ and $f''(x) \geq 0$. Define the following infinite sequence recursively:*

$$\begin{aligned} x_0 &= c \\ x_{n+1} &= x_n - \frac{f(x_n}{f'(x_n)} \end{aligned}$$

Then the sequence x_n is monotone decreasing and it has a limit of a.

Reason: Our hypotheses guarantee that $x_{n+1} \leq x_n$. Now we will show by induction that $x_n \geq a$. We assumed that $x_0 > a$. Now, suppose that $x_n \geq a$. We want to show that

$$a \leq x_n - \frac{f(x_n)}{f'(x_n)}.$$

This is equivalent to showing

$$\begin{aligned} f'(x_n)(a - x_n) &\leq -f(x_n) \\ f(x_n) + f'(x_n)(a - x_n) &\leq 0 \\ -f(x_n) - f'(x_n)(a - x_n) &\geq 0 \end{aligned}$$

Since $f(a)=0$ this last inequality is equivalent to

$$f(a)-f(x_n)-f'(x_n)(a-x_n)\geq 0$$

If we apply our error estimate for the linear approximation we know that there is some $z\in(a,x_n)$ so that

$$f(a)-f(x_n)-f'(x_n)(a-x_n)=\frac{1}{2}f''(z)(a-x_n)\geq 0$$

which is what we wanted to do. Now we know that

$$\lim_{n\to\infty}x_n=L\geq a.$$

We need to show that $L=a$. Suppose $L>a$. We know that

$$f'(x_n)(x_{n+1}-x_n)=f(x_n).$$

Take the limit of both sides of this equation as n tends to infinity. Since f and f' are continuous on $[L,c]$ we have $0=f'(L)(L-L)=f(L)$, which is impossible since f is strictly increasing on $[a,c]$. This means that $L=a$. **QED**

For example, suppose that we want to solve $x^2-5=0$ for $x>0$. We let $a=\sqrt{5}$ and $b=10$. Take $c=3$. Then since $f'(x)=2x$ we have

$$\begin{aligned}x_0&=3\\ x_{n+1}&=x_n-\frac{x_n^2-5}{2x_n}=\frac{x_n^2+5}{2x_n}\end{aligned}$$

1.3 Linear Interpolation

Recall that if (a,b) and (c,d) are distinct points not lying on a vertical line then an equation of the line passing through them is

$$y=b\frac{x-c}{a-c}+d\frac{x-a}{c-a}. \qquad (1)$$

If we have a function f defined on an interval I, and a and c lie in I, we may define a function $C(x)$ by

$$C(x)=f(a)\frac{x-c}{a-c}+f(c)\frac{x-a}{c-a}. \qquad (2)$$

This function $C(x)$ is called the linear interpolation of f on the interval $[a,c]$. (We are supposing $a<c$.)

The graph of $y=C(x)$ is the graph of line containing the chord to the graph of $y=f(x)$ with endpoints $(a,f(a))$ and $(c,f(c))$. A question then arises concerning the relation between the values $f(x)$ and $C(x)$ for $a<x<c$. The second Rolle's Theorem supplies an answer.

Theorem 4 (Linear Interpolation) *Suppose that f is continuous on $[a,c]$ and $f''(x)$ exists for all $a<x<c$. Define $C(x)$ as in (2). Then for each x in (a,c) there is some $d\in(a,c)$ so that*

$$f(x)-C(x)=-\frac{1}{2}f''(d)(x-a)(c-x). \qquad (3)$$