320 Fall 2021, Tuesday, Oct 26, 2021

Section 6.4: Taylor's Theorem

1. Derivatives of order greater than 1: $f'', f''' = f^{(3)}, \dots$
2. **Defn:** Taylor polynomial of f at x_0:

$$P_n(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 \cdots + \frac{f^{(n)}(x_0)}{n}(x - x_0)^n.$$

3. **Taylor's Theorem:** Let $n \in \mathbb{N}$, $I = [a, b]$, $f : I \to \mathbb{R}$ be such that f and its derivatives up to order n are continuous on I and that $f^{(n=1)}$ exists on (a, b). If $x_0 \in I$ then for any $x \in I$ there is a point c between x_0 and x so that

$$f(x) = P_n(x) + \frac{f^{(n+1)}(c)}{(n+1)}(x - x_0)^{n+1} = P_n(x) + R_n(x).$$

Proof: Fix x and let J be the closed interval with endpoints x and x_0. Define

$$F(t) = f(x) - f(t) - (x - t)f'(t) - \cdots - \frac{(x - t)^n}{n} f^{(n)}(t).$$

Note $F(x) = 0$. Differentiate with respect to t and use product rule :

$$F'(t) = 0 - f'(t) - f'(t) - (x - t)f''(t) - \cdots - \frac{(x - t)^{n-1}}{(n-1)} f^{(n)}(t) - \frac{(x - t)^n}{n} f^{(n+1)}(t).$$

All the terms cancel except the last, so

$$F'(t) = -\frac{(x - t)^n}{n} f^{(n+1)}(t).$$

Define G on J by

$$G(t) = F(t) - \left(\frac{(x - t)}{x - x_0}\right)^{n+1} F(x_0).$$

Then $G(x) = G(x_0) = 0$.

By Rolle's theorem there is a $c \in J$ so that

$$0 = G'(c) = F'(c) = (n + 1) - \frac{(x - c)^n}{(x - x_0)^{n+1}} F(x_0).$$

Thus

$$\begin{aligned} F(x_0) &= \frac{-1}{n+1} - \frac{(x - x_0)^{n+1}}{(x - c)^n} F'(c) \\ &= \frac{1}{n+1} - \frac{(x - x_0)^{n+1}(x - c)^n}{(x - c)^n n} f^{(n+1)}(c) \\ &= -\frac{f^{(n+1)}(c)}{(n+1)}(x - x_0)^{n+1}. \quad \square \end{aligned}$$

4. $e = 1 + \frac{1}{2} + \frac{1}{6} + \cdots + \frac{1}{n} + \ldots$

We assume the derivative of e^x is itself. Then the Taylor polynomial for e^x at $x_0 = 0$ is

$$1 + x + \frac{1}{2}x^2 + \ldots \frac{1}{n}x^n$$

and the remainder is

$$\frac{f^{(n+1)}(c)}{(n+1)}(x - x_0)^{n+1} = e^c(n+1)x^{n+1}$$

for some $c \in [0, x]$. Taking $x = 1$ we see $R_n(1) \to 0$, so

$$e = e^1 = \lim_{n\to\infty}(1 + \frac{1}{2} + \frac{1}{6} + \cdots + \frac{1}{n}).$$

5. $1 - x^2 \leq \cos x$ for all $x \in \mathbb{R}$.

By Taylor' theorem

$$\cos(x) = 1 - x^2/2 + R_2(x)$$

and

$$R_2(x) = \frac{1}{6}f'''(c)x^3 = \frac{1}{6}\sin(c)x^3.$$

If $0 \leq c \leq x \leq \pi$ then $R_2(x) \geq 0$. Similarly if $-\pi \leq x \leq c \leq 0$.

For $|x| \geq \pi$, $1 - x^2 < -3 < \cos(x)$, so this is trivially true.

6. **Theorem 6.4.4:** Suppose I is an interval, x_0 is an interior point of I and $n \geq 2$. Suppose $f', \ldots, f^{(n)}$ are continuous on a neighborhood of x_0 and the first $n-1$ derivatives vanish at x_0.

(i) If n is even and $f^{(n)}(x_0) > 0$, then f has a relative maximum at x_0.

(ii) If n is even and $f^{(n)}(x_0) < 0$, then f has a relative minimum at x_0.

(iii) If n is odd, then f has neither a relative minimum or maximum at x_0.

Proof: see text.

7. **Defn:** A function $f : I \to \mathbb{R}$ is convex if for any $t \in [0, 1]$ and $x, y \in I$, we have

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y).$$

8. **Facts:** convex functions are continuous and left and right derivatives exist at every point.

9. **Theorem 6.4.6:** If f has a second derivative on I then f is convex iff $f''(x) \geq 0$ for all $x \in I$.

Proof:

Suppose f is convex. An exercise shows that

$$f''(a) = \lim_{h\to 0}\frac{f(a+h) - 2f(a) + f(a-h)}{h^2}.$$

If f is convex the quotient is ≥ 0 so $f''(a) \geq 0$.

To prove the converse, use Taylor's theorem with $a_1, x_2 \in I$ and $x_0 = (1-t)x_1 + tx_2$ to get

$$f(x) = f(x_0) + f'(x_0)(x_1 - x_0) + \frac{1}{2} f''(c_1)(x_1 - x_0)^2$$

for some c_1 between x_0 and x_1. Similarly,

$$f(x) = f(x_0) + f'(x_0)(x_2 - x_0) + \frac{1}{2} f''(c_2)(x_2 - x_0)^2$$

for some c_2 between x_0 and x_2. If $f'' \geq 0$ then the remainder terms are positive (remember the square), so

$$\begin{aligned}
(1-t)f(x_1) + tf(x_2) &= f(x_0) + f'(x_0)((1-t)x_1 + tx_2 - x_0) \\
&\quad + \frac{1}{2}(1-t)f''(c_1)(x_1 - x_0)^2 \\
&\quad + \frac{1}{2} t f''(c_2)(x_2 - x_0)^2 \\
&= f(x_0) + R \\
&\geq f(x_0) = f((1-t)x_1 + tx_2). \quad \square
\end{aligned}$$

10. **Newton's method:** Let $I = [a, b]$ and let f be twice differentiable on I. Suppose $f(a) < 0 < f(b)$. Suppose there are constants m, M so that

$$0 < m \leq |f'(x)| \text{ and } |f''(x)| \leq M < \infty$$

on I. Them there exists a subinterval J of I containing a zero r of f such that for any $x \in J$ the sequence defined by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

belongs to J and (x_n) converges to r. Moreover

$$|x_n - r| \leq \frac{M}{2m} |x_n - r|^2$$

for all $n \in \mathbb{N}$.

Proof: see text.

11. Newton's method need not converge if we start too far from root. Iterates can become periodic or chaotic.

Section 7.1: Riemann Integral

1. **Defn:** If $I = [a, b]$ a partition of I is finite set $\mathcal{P}$

$$a = x_0 < x_1 < \cdots < x_n = b.$$

These divide I into disjoint (except for endpoints) intervals $I_k = [x_{k-1}, x_k]$ for $k = 1, \ldots, n$. The norm of $\mathcal{P}$ is

$$\|\mathcal{P}\| = \max_k |x_k - x_{k-1}|.$$

2. **Defn:** A tag is a choice of a point $t_k \in I_k$. A tagged partition is a partition along with a tag for each interval.
4. Given a tagged partition, the associated Riemann sum for a function $f : I \to \mathbb{R}$ is

$$S(f, \mathcal{P}) = \sum_{k=1}^{n} f(t_k)(x_k - x_{k-1}).$$

5. **Defn:** A function $f : I \to \mathbb{R}$ is Riemann integrable on $[a, b]$ if there is a $L \in \mathbb{R}$ so that for every $\epsilon > 0$ there is a δ so that $\|\mathcal{P}\| < \delta$ implies $|S(f, \mathcal{P}) - L| < \epsilon$.
   In this case we write $f \in \mathcal{R}[a, b]$.
6. If L exists we usually write it as

$$\int_a^b f(x)dx.$$

7. **Theorem 7.1.2:** If $f \in \mathcal{R}[a, b]$ then the value of the integral is uniquely determined.
Proof: sketch or see text.
8. **Theorem 7.1.3:** If $g \in \mathcal{R}[a, b]$ and $f = g$ except at finitely many points, then $f \in \mathcal{R}[a, n]$ and $\int f = \int g$.
Proof: If suffices to prove this for one point $c \in [a, b]$, and then apply induction.

Let $L = \int_a^b g$ and $f = g$ except at c. The for any tagged partition, the Riemann sums for f and g are the same except possible for intervals with tag c, and there are at most two such intervals. Thus

$$|S(f, \mathcal{P}) - S(g, \mathcal{P})| \leq 2(|g(c)| + |f(c)|)\|\mathcal{P}\|.$$

Given $\epsilon > 0$ choose

$$\delta_1 \leq \frac{\epsilon}{4(|f(c)| + |g(c)|)},$$

and δ_2 so that $\|\mathcal{P}\| < \delta_2$ implies

$$|L - S(g, \mathcal{P})| < \epsilon/2.$$

Take $\delta = \min(\delta_1, \delta_2)$. Then if $\|\mathcal{P}\| < \delta$ we get

$$|L - S(f, \mathcal{P})| \leq |L - S(g, \mathcal{P})| + |S(f, \mathcal{P}) - S(g, \mathcal{P})| < \epsilon/2 + \epsilon/2. \quad \square$$

9. Examples:

constant functions
step functions
continuous functions
Thomae's function
Any function with a countable number of discontinuities
Any function with a zero length set of discontinuities

10. Non-examples
Dirichlet function

11. **Theorem 7.1.5:** If $f, g \in \mathcal{R}[a,b]$ and $k \in \mathbb{R}$, then
(a) $kf \in \mathcal{R}[a,b]$ and $\int_a^b kf = k\int_a^b f$.
() $f + g \in \mathcal{R}[a,b]$ and $\int_a^b (f-g) = \int_a^b f + \int_a^b g$.
(a) if $f \le g$ then $\int_a^b f \le \int_a^b g$.

Proof: It is easy to check from the definitions that

$$S(kf, \mathcal{P}) = kS(f, \mathcal{P}),$$

$$S(f+g, \mathcal{P}) = S(f, \mathcal{P}) + S(g, \mathcal{P}),$$

$$S(f, \mathcal{P}) \le S(g, \mathcal{P}),$$

and from these it is easy to deduce the stated results.

To prove (b), note that

$$|S(f+g, \mathcal{P}) - \int f - \int g| \le |S(f, \mathcal{P}) - \int f| + |S(g, \mathcal{P}) - \int g|$$

and the two terms on the right are less than $\epsilon/2$ if $\|\mathcal{P}\| < \delta$ is small enough.

To prove (c), choose $\|\mathcal{P}\| < \delta$ small enough so that

$$\int f - \epsilon/2 \le S(f, \mathcal{P})$$

$$\int g + \epsilon/2 \ge S(f, \mathcal{P})$$

so then

$$\int f \le \epsilon + \int g,$$

for all $\epsilon > 0$. This implies $\int f \le \int g$. □

12. **Theorem 7.1.6:** If $f \in \mathcal{R}[a,b]$ then f is bounded on $[a,b]$.

Proof: Suppose f is integrable with $\int f = L$. Then taking $\epsilon = 1$, we can choose $\delta > 0$ so that $\|\mathcal{P}\| < \delta$ implies

$$|S(f, \mathcal{P})| \le |L| + 1.$$

If f is not bounded on I then it is also not bounded on at least one of the subintervals I_k. Thus we can choose the tag in I_k so that $f(t_k)(x_k - x_{k-1})$ is as large as we wish, say

$$f(t_k)(x_k - x_{k-1}) > |L| + 1 + \| \sum_{j \neq k} f(t_j)(x_j - x_{j-1}).$$

This contradicts the triangle inequality. $\square$

13. Thomae's function f is Riemann integrable on $[0, 1]$.

For any $\epsilon > 0$ there are only finitely many points M_ϵ where $f(x) > \epsilon/2$, so at most $2M_\epsilon$ tags where this happens. Thus $\|\mathcal{P}\| < \delta$ implies

$$S(f, \mathcal{P}) \leq 1 \cdot M_\epsilon \cdot \delta + \epsilon/2 \sum (x_k - x_{k-1}) \leq \delta M_\epsilon + \epsilon.$$

Choose $\delta \leq \epsilon/(2M_\epsilon)$.Then

$$S(f, \mathcal{P}) < \epsilon/2 + \epsilon/2 = \epsilon.$$

Thus Thomae's function is Riemann integrable with $\int f = 0$.