

Taylor's Theorem

1-D: $f(x_0 + \Delta x) = \sum_{k=0}^{r-1} \frac{1}{k!} \Delta x^k \partial_x^k f(x_0)$
 $+ \frac{1}{r!} \Delta x^r \partial_x^r f(x_0 + t \Delta x)$
 where $0 \leq t \leq 1$. [Can't depend continuously on Δx and f^2 differentiability]

(+) \mathbb{R}^n : $f(x_0 + \Delta x) = \sum_{s=0}^{r-1} \frac{1}{s!} \sum_{i \in \mathbb{Z}_n^s} \Delta x_{i_1} \dots \Delta x_{i_s} \partial_{i_1} \dots \partial_{i_s} f(x_0)$
 $+ \frac{1}{r!} \sum_{i \in \mathbb{Z}_n^r} \Delta x_{i_1} \dots \Delta x_{i_r} \partial_{i_1} \dots \partial_{i_r} f(x_0 + t \Delta x)$
 where $0 \leq t \leq 1$. ($\partial_x \equiv \partial_{x_1}$)

Pf. Let $F(t) = f(x_0 + t \Delta x)$
 Apply the chain rule and the 1-D Taylor thm:
 $F'(t) = \sum_i \Delta x_i \partial_{x_i} f(x_0 + t \Delta x)$
 $F^{(s)}(t) = \sum_{i \in \mathbb{Z}_n^s} \Delta x_{i_1} \dots \Delta x_{i_s} \partial_{x_{i_1}} \dots \partial_{x_{i_s}} f(x_0 + t \Delta x)$
 etc.

(*) \mathbb{R}^n (more popular version)
 $f(x_0 + \Delta x) = \sum_{s=0}^{r-1} \sum_{k \in \mathbb{I}_n^s} \frac{1}{k!} \Delta x^k \partial_x^k f(x_0)$
 $+ \sum_{k \in \mathbb{I}_n^r} \frac{1}{k!} \Delta x^k \partial_x^k f(x_0 + t \Delta x)$
 where $0 \leq t \leq 1$

Where:
 $\mathbb{I}_n^s \equiv \{k: \mathbb{Z}_n \rightarrow \mathbb{Z}^+ \mid \sum_{i=1}^n k_i = s\}$
 where $\mathbb{Z}^+ =$ nonnegative integers
 $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$
 $k! \equiv \prod_{i=1}^n k_i! = k_1! \dots k_n!$
 $\Delta x^k \equiv \Delta x_1^{k_1} \dots \Delta x_n^{k_n}$
 $\partial_x^k \equiv \partial_{x_1}^{k_1} \dots \partial_{x_n}^{k_n}$

example: 2-D

$f(a_1 + \Delta x_1, a_2 + \Delta x_2)$
 $= \sum_{s=0}^{r-1} \sum_{(k_1, k_2): k_1+k_2=s} \frac{1}{k_1! k_2!} \Delta x_1^{k_1} \Delta x_2^{k_2} \partial_{x_1}^{k_1} \partial_{x_2}^{k_2} f(a)$
 $+ \sum_{(k_1, k_2): k_1+k_2=r} \text{[same stuff]} f(a + t \Delta x)$

Pf. Need to see that the s^{th} term is the same for both Taylor sums.
 To see this, apply $\partial_{x_1}^{k_1} \dots \partial_{x_n}^{k_n}$ to the sum and observe that you get $\partial_{x_1}^{k_1} \dots \partial_{x_n}^{k_n} f(a)$ at a .
 Suffices to show that each term is correct (i.e. has correct coefficient).

Note:
 One can also verify the first Taylor series this way: Apply $\partial_{j_1} \dots \partial_{j_s}$ & eval. at $\Delta x = 0$.
 Observe that the only terms which survive are those for which a permutation of the sequence of partials $\partial_{j_1} \dots \partial_{j_s}$ can be matched up to the $\Delta x_{i_1} \dots \Delta x_{i_s}$.
 These all have the same coefficient $\partial_{j_1} \dots \partial_{j_s} f(x_0 + t \Delta x)$, since the partials can also be matched up, and partials commute. So it suffices to show that:
 $\partial_{j_1} \dots \partial_{j_s} \sum_{i \in \mathbb{Z}_n^s} \Delta x_{i_1} \dots \Delta x_{i_s} = s!$

But LHS = $\sum_{i \in \mathbb{Z}_n^s} \partial_{j_1} \dots \partial_{j_s} \Delta x_{i_1} \dots \Delta x_{i_s}$
 $= \sum_{i \in \mathbb{Z}_n^s} \sum_{k \in P_s} (\partial_{j_{k_1}} \Delta x_{i_{k_1}}) \dots (\partial_{j_{k_s}} \Delta x_{i_{k_s}})$
 where $P_s =$ set of all permutations of \mathbb{Z}_s . (comes from repeated application of product rule.)
 (summing over all ways of matching differential operators to a particular choice of Δx 's.)
 $= \sum_{i \in \mathbb{Z}_n^s} \sum_{k \in P_s} \delta_{j_{k_1} i_{k_1}} \dots \delta_{j_{k_s} i_{k_s}}$
 where $\delta_x = \begin{cases} 1 & \text{if } x=0 \\ 0 & \text{else} \end{cases}$
 $= \sum_{k \in P_s} 1$ (There is only one i that gives nonzero for a given k)
 $= s!$, as desired.

Another argument for equivalence of (+) and (*):

Pick arbitrary term k , \underline{k} from (*).
 The number of times this term appears in (+) is $\binom{s}{\underline{k}} = s$ -choose- $\underline{k} = \frac{s!}{k!}$
 (i.e. the number of ways of choosing the s indices i_1, \dots, i_s so that Δx_{i_l} appears k_l times for each l).

Vector notation for (+):

(+) $f(x_0 + \Delta x) = \sum_{s=0}^{r-1} \frac{1}{s!} (\Delta x \cdot \nabla)^s f(x_0)$
 $+ \frac{1}{r!} (\Delta x \cdot \nabla)^r f(x_0 + t \Delta x)$
 where $0 \leq t \leq 1$.

Operator notation for (+):

(+) $f(x_0 + \Delta x) = \sum_{s=0}^{r-1} \frac{1}{s!} \left(\sum_{i=1}^n \Delta x_i \partial_{x_i} \right)^s f(x_0)$
 $+ \frac{1}{r!} \left(\sum_{i=1}^n \Delta x_i \partial_{x_i} \right)^r f(x_0 + t \Delta x)$
 where $0 \leq t \leq 1$

Proof of 1-D Taylor:

WLOG $x_0 = 0$.

$$f(h) = f(0) + \int_0^h f'(x) dx \leftarrow R_0(h)$$

$$R_0(h) = \int_0^h [(x-h)f']' - (x-h)f'' dx$$

$$= \underbrace{[(x-h)f']_0^h}_h f'(0) + \underbrace{\int_0^h (h-x)f'' dx}_{R_1(h)}$$

$$R_1(h) = \int_0^h \left[\frac{(h-x)^2}{-2} f'' \right]' - \frac{(h-x)^2}{-2} f''' dx$$

$$= \underbrace{\left[\frac{(h-x)^2}{2} f'' \right]_0^h}_{\frac{1}{2} h^2 f''(0)} + \underbrace{\int_0^h \frac{1}{2} (h-x)^2 f''' dx}_{R_2(h)}$$

Claim

$$R_n(h) = \int_0^h \frac{1}{n!} (h-x)^n f^{(n+1)} dx$$

Inductive verification:

$$R_n(h) = \int_0^h \left[\frac{1}{(n+1)!} (h-x)^{n+1} f^{(n+1)} \right]' + \frac{(h-x)^{n+1}}{(n+1)!} f^{(n+2)} dx$$

$$= \frac{h^{n+1}}{(n+1)!} f^{(n+1)}(0) + R_{n+1}(h) \quad \checkmark$$

Sampling form of remainder

$$R_n(h) = \int_0^h \frac{1}{n!} (h-x)^n f^{(n+1)}(x) dx$$

Recall weighted mean value theorem:
 Suppose $g \geq 0$ on Ω ,
 and f & g cont. on Ω .
 Then $\int_{\Omega} fg = f(p) \int_{\Omega} g$
 for some $p \in \Omega$.

$$\text{So } R_n(h) = f^{(n+1)}(p) \int_0^h \frac{1}{n!} (h-x)^n dx$$

for some $p \in [0, h]$

$$\text{So } R_n(h) = f^{(n+1)}(p) \frac{h^{n+1}}{(n+1)!} \quad \checkmark$$

Can p depend continuously on h and f ?

1-D Taylor

$$f(h) = f(0) + hf'(0) + \frac{1}{2} h^2 f''(0) + \dots + \frac{1}{n!} h^n f^{(n)}(0) + R_n(h)$$

where

$$R_n(h) = \int_0^h \frac{1}{n!} (h-x)^n f^{(n+1)}(x) dx$$

$$= \frac{1}{(n+1)!} h^{n+1} f^{(n+1)}(p)$$

where $p \in [0, h]$.

Cor

$$f(x_0 + \Delta x) = \left[\sum_{k=0}^n \frac{1}{k!} \Delta x^k f^{(k)}(x_0) \right] + R_n(\Delta x)$$

where

$$R_n(\Delta x) = \int_0^{\Delta x} \frac{1}{n!} (\Delta x - \tilde{\Delta x})^n f^{(n+1)}(x_0 + \tilde{\Delta x}) d\tilde{\Delta x}$$

$$= \frac{1}{(n+1)!} \Delta x^{n+1} f^{(n+1)}(x_0 + p)$$

for some $p \in [0, \Delta x]$.

Mean Value Theorem (See Guran, Differential & Integral Calculus, Volume 1.)

Let Ω be a connected, compact set.

Let $f: \Omega \rightarrow \mathbb{R}$ be continuous

Let $g: \Omega \rightarrow \mathbb{R}, g \geq 0$, be continuous (integrable enough?)

Then $\int_{\Omega} fg = f(\xi) \int_{\Omega} g$ for some $\xi \in \Omega$.

Pf Let $m = \min_{\Omega} f$.

Let $M = \max_{\Omega} f$.

Since $g \geq 0$, $mg(x) \leq f(x)g(x) \leq Mg(x) \quad (\forall x \in \Omega)$

So $m \int_{\Omega} g(x) dx \leq \int_{\Omega} f(x)g(x) dx \leq M \int_{\Omega} g(x) dx$.

So $\int_{\Omega} f(x)g(x) dx = \mu \int_{\Omega} g(x)$, where $m \leq \mu \leq M$.

$f(\Omega)$ is connected and compact (since f is continuous).

So f attains its minimum & maximum on Ω .

Let $f(x_1) = m$ and $f(x_2) = M$.

Let $\tau: [0, 1] \rightarrow \Omega$ be a continuous path

from x_1 to x_2 . (So $\tau(x_1) = 0$ and $\tau(x_2) = 1$.)

Note $f \circ \tau: [0, 1] \rightarrow [m, M]$ is continuous & so onto.

So $\exists t \in [0, 1]$ s.t. $f \circ \tau(t) = \mu$.

So $\exists \xi = \tau(t) \in \Omega$ s.t. $f(\xi) = \mu$.

So $\int_{\Omega} fg = f(\xi) \int_{\Omega} g$, as desired.

Can ξ depend continuously on Ω, f , and g ?

Taylor's Theorem (For sharp proofs)

Let $f(x) \in \mathcal{C}^r$, $x \in \mathbb{R}^n$. Write:

$$f(x_0 + \Delta x) = \sum_{s=0}^r \frac{1}{s!} \sum_{i \in \mathbb{Z}_n^s} \Delta x_{i_1} \Delta x_{i_2} \dots \Delta x_{i_s} \partial_{i_1} \dots \partial_{i_s} f(x_0) + \text{error}(\Delta x)$$

Then: $\lim_{|\Delta x| \rightarrow 0} \frac{\text{error}(\Delta x)}{|\Delta x|^r} = 0$

Proof i.e. $\text{error}(\Delta x) = o(|\Delta x|^r)$
 Let $\tilde{f}_s(x_0, \Delta x) = \sum_{i \in \mathbb{Z}_n^s} \Delta x_{i_1} \dots \Delta x_{i_s} \partial_{i_1} \dots \partial_{i_s} f(x_0)$

Let $\tilde{f}(x_0, \Delta x) = \sum_{s=0}^r \frac{1}{s!} \tilde{f}_s(x_0, \Delta x)$

So $\text{error}(\Delta x) = f(x_0 + \Delta x) - \tilde{f}(x_0, \Delta x)$

Use Taylor's theorem to write:

$$f(x_0 + \Delta x) = \sum_{s=0}^{r-1} \frac{1}{s!} \tilde{f}_s(x_0, \Delta x) + \frac{1}{r!} \tilde{f}_r(\xi, \Delta x)$$

where $\xi \in [x_0, x_0 + \Delta x]$.

$$f(x_0 + \Delta x) = \sum_{s=0}^r \frac{1}{s!} \tilde{f}_s(x_0, \Delta x) + \underbrace{\frac{1}{r!} (\tilde{f}_r(\xi, \Delta x) - \tilde{f}_r(x_0, \Delta x))}_{\text{error}(\Delta x)}$$

So $\text{error}(\Delta x) = \frac{1}{r!} \sum_{i \in \mathbb{Z}_n^r} \Delta x_{i_1} \dots \Delta x_{i_r} \left[\partial_{i_1} \dots \partial_{i_r} f(\xi) - \partial_{i_1} \dots \partial_{i_r} f(x_0) \right]$
 Call $A_{i_1 \dots i_r}(\Delta x)$.

Since $\text{error}(\Delta x) = f(x_0 + \Delta x) - \sum_{s=0}^r \frac{1}{s!} \tilde{f}_s(x_0, \Delta x)$,
 error is continuous on the compact set $[0, \Delta x]$, so it is bounded.

$A_{i_1 \dots i_r}(\Delta x)$ is bounded in any compact domain $|\Delta x| < \delta$, because $f \in \mathcal{C}^r$.

Also, $A_{i_1 \dots i_r}(\Delta x) \rightarrow 0$ as $\Delta x \rightarrow 0$, since $\partial_{i_1} \dots \partial_{i_r} f(\xi) \rightarrow \partial_{i_1} \dots \partial_{i_r} f(x_0)$ as $\Delta x \rightarrow 0$

So $\frac{\text{error}(\Delta x)}{|\Delta x|^r} = \frac{1}{r!} \sum_{i \in \mathbb{Z}^r} \hat{\Delta x}_{i_1} \dots \hat{\Delta x}_{i_r} A_{i_1 \dots i_r}(\Delta x)$
 (where $\hat{\Delta x}_i = \frac{\Delta x_i}{|\Delta x|}$)

So $\left| \frac{\text{error}(\Delta x)}{|\Delta x|^r} \right| \leq \frac{1}{r!} \sum_{i \in \mathbb{Z}^r} A_{i_1 \dots i_r}(\Delta x) \leq \frac{1}{r!} \sum n^r \max |A_{i_1 \dots i_r}(\Delta x)| \rightarrow 0$ as $\Delta x \rightarrow 0$.

Special cases

\square^1 Let $f \in \mathcal{C}^1$.

Then $f(x_0 + \Delta x) = f(x_0) + \Delta x_i \partial_i f(x_0) + \text{error}(\Delta x)$

where $\lim_{|\Delta x| \rightarrow 0} \frac{\text{error}(\Delta x)}{|\Delta x|} = 0$

i.e. $\text{error}(\Delta x) = o(|\Delta x|)$

\square^2 Let $f \in \mathcal{C}^2$

Then $f(x_0 + \Delta x) = f(x_0) + \Delta x_i \partial_i f(x_0) + \Delta x_i \Delta x_j \partial_i \partial_j f(x_0) + \text{error}(\Delta x)$

where $\lim_{|\Delta x| \rightarrow 0} \frac{\text{error}(\Delta x)}{|\Delta x|^2} = 0$

i.e. $\text{error}(\Delta x) = o(|\Delta x|^2)$