Multivariate Calculus 1 - MA Math Camp 2022

Andrea Ciccarone

August 25, 2022

Derivatives in one dimension

- The fundamental concept in calculus is the **derivative**. The main idea underlying derivatives is a rate of change.
- Consider a function $f : \mathbb{R} \to \mathbb{R}$ and take:

$$\frac{f(x) - f(x_0)}{x - x_0}$$

  This is giving us the average rate of change of the function f between x and x_0.
- The idea of the derivative is to let x go to x_0 and study what happens at this rate of change. In this sense, we can talk about the **instantaneous rate of change**:

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

Definition of Derivative

Definition 1.1

- Let $A \subseteq \mathbb{R}$, and $x_0 \in A \cap A'$. A function $f : A \to \mathbb{R}$ is said to be **differentiable** at x_0 iff the limit

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

  exists. In that case, In that case, define the **derivative of** f **at** x_0 as the limit above, denoted as $f'(x_0)$.
- A function $f : A \to \mathbb{R}$ is said to be **differentiable** iff $A \subseteq A'$ and f is differentiable at any $x_0 \in A$.
- Let $\hat{A}$ be the set of points in $A \cap A'$ at which f is differentiable. Then the function $f' : \hat{A} \to \mathbb{R}$ is called the **derivative (function)** of f.

An Example

- Notice that by just applying the change of variable $x = x_0 + h$ we can rewrite the limit as:

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

- For example, if we want to show that the function $f : \mathbb{R} \to \mathbb{R}$ s.t. $f(x) = \sqrt{x}$ is differentiable (and compute the associated derivative) we observe:

$$\frac{\sqrt{x+h} - \sqrt{x}}{h} = \frac{\sqrt{x+h} - \sqrt{x}}{x + h - x} = \frac{1}{\sqrt{x+h} + \sqrt{x}}$$

so that:

$$\lim_{h \to 0} \frac{\sqrt{x+h} - \sqrt{x}}{h} = \frac{1}{2\sqrt{x}}$$

Differentiability implies Continuity

- If a function is differentiable at x_0 it is continuous at x_0. The contrary is not always true (Example?)
- Remember that a function is continuous at x_0 iff the limit of the function at x_0 is equal to the value of the function at x_0
- We can thus prove that a differentiable function is continuous by writing:

$$
\begin{aligned}
\left[\lim_{x\to x_0} f(x)\right] - f(x_0) &= \lim_{x\to x_0}\left[\frac{f(x)-f(x_0)}{x-x_0}\cdot(x-x_0)\right] \\
&= \lim_{x\to x_0}\left[\frac{f(x)-f(x_0)}{x-x_0}\right]\cdot\lim_{x\to x_0}(x-x_0) \\
&= f'(x)\cdot 0 = 0
\end{aligned}
$$

Common Derivatives and Properties

- **Common Derivatives:**
  - $(x^\alpha)' = \alpha x^{\alpha-1}$
  - $(\ln x)' = \frac{1}{x}$
  - $(e^x)' = e^x$
  - $(\sin x)' = \cos x$
- **Properties:**
  - $(f+g)' = f' + g'$
  - $(\lambda f)' = \lambda f'$
  - $(fg)' = f'g + fg'$
  - $(f/g)' = \frac{f'g - fg'}{g^2}$

Derivatives as Affine Transformations

- An other interpretation of a derivative is that it is the best linear approximation of a function at x_0.
- In order to appreciate this interpretation, let us introduce first order expansions

Definition 1.4

Let $f : A \to \mathbb{R}$ and $x_0 \in A \cap A'$. We say that f admits a first order expansion around x_0 if there exists $a, b \in \mathbb{R}$ and a function $\epsilon : A \to \mathbb{R}$ such that:

$$\forall x \in A, f(x) = a + b(x - x_0) + (x - x_0)\varepsilon(x)$$
$$\text{and } \lim_{x \to x_0} \varepsilon(x) = 0$$

Derivatives as Affine Transformations (ctd.)

- If a function is differentiable at x_0 it will always have a first order expansion at x_0 (set $\epsilon(x) = \frac{f(x)-f(x_0)}{x-x_0} - f'(x_0)$, $a = f(x_0)$, $b = f'(x_0)$)
- If a function has a first order expansion, it will be differentiable and we can show $a = f(x_0)$, $b = f'(x_0)$:
  - If $\forall x \in A, f(x) = a + b(x - x_0) + (x - x_0)\varepsilon(x)$ then $f(x_0) = a$
  - For $x \in A \setminus \{x_0\}$ we have:

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{f(x) - a}{x - x_0} = b + \epsilon(x) \xrightarrow[x \to x_0]{} b$$

  - Thus $f'(x_0) = b$
- All this combined implies the next result.

Derivatives as Affine Transformations (ctd.)

Theorem 1.5

Let $f : A \to \mathbb{R}$ and $x_0 \in A \cap A'$. The following are equivalent :

1. f is differentiable at x_0
2. f has a first order expansion at x_0

Furthermore the coefficients of the first order expansion when they exist are $a = f(x_0)$, $b = f'(x_0)$.

- The function $f(x_0) + (x - x_0)f'(x_0)$ is the affine transformation of f at x_0. Geometrically speaking, it is line that is tangent to f at x_0.

- The following theorem is an important result with many useful implications:

Mean Value Theorem

Let $f : [a, b] \to \mathbb{R}$, differentiable on (a, b), and continuous on $[a, b]$. Then there exists $x \in (a, b)$ s.t.

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

- In words, MVT states that we can find some x in the interior of the interval $[a, b]$ such that the average rate of change is equal to the instantaneous rate of change at that point
- We are not going to prove it, but the proof is on the lecture notes (you are also going to see it in Math Methods)

MVT and IVT

- An important implication of MVT is that if $f' > 0$ then f is stringly increasing.
- Take any x_1, x_2 in (a, b) with $x_1 < x_2$. By MVT there exists some $x \in (x_1, x_2)$ s.t.

$$f'(x) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

- Thus $f(x_2) - f(x_1) = f'(x) \cdot (x_2 - x1) > 0 \implies f$ is strictly increasing.

Intermediate Value Theorem

Let $f : [a, b] \to \mathbb{R}$ continuous and u is a number between $f(a)$ and $f(b)$, then there exists $c \in [a, b]$ s.t. $u = f(c)$.

L'Hospital Rule

- Using MVT, it is also possible to obtain a result which is very useful in computing some limits.
- In order to state this result we need to define the extended real line $\bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}$ by extending the order $\leq$ s.t. $+\infty > a$ and $-\infty < a$ for any $a \in \mathbb{R}$.

L'Hospital Rule

Let $-\infty \leq a < b \leq +\infty$, and $f : (a, b) \to \mathbb{R}$ and $g : (a, b) \to \mathbb{R} \setminus \{0\}$ are differentiable in (a, b). If $\lim_{x \to a} f(x)$ and $\lim_{x \to a} g(x)$ are both 0 or $\pm\infty$, and $\lim_{x \to a} f'(x)/g'(x)$ has a finite value or is $\pm\infty$, then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}$$

The statement is also true for $x \to b$.

L'Hospital Rule - Example

- Suppose we want to find the limit of $\frac{\ln(x)}{\sqrt{x}}$ when $x \to +\infty$.
- Notice that both the nominator and the denominator diverge to $+\infty$, so we can apply L'Hospital
- In particular, we have:

$$\frac{(\ln x)'}{(\sqrt{x})'} = \frac{\frac{1}{x}}{\frac{1}{2\sqrt{x}}} = \frac{2}{\sqrt{x}} \to 0$$

- We can thus conclude that our function converges to 0 as x diverges to $+\infty$.

Total Derivatives - Introduction

- As economists, you are going to work a lot with functions from $\mathbb{R}^n$ to $\mathbb{R}$ (eg. utility functions).
- We thus want to extend the concept of derivatives to multivariate functions.
- We mentioned that the derivative at x_0 is slope of the "best" linear approximation we can find for $f(x)$ around x_0:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0)$$

- "Best" means that the relative error term $\epsilon(x)$ goes to 0. In other words, $f'(x_0)$ is the value of m such that:

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - m(x - x_0)}{x - x_0} = 0$$

12

Total Derivatives

- If $f : \mathbb{R}^n \to \mathbb{R}^m$ then we can write a linear approximation (around x_0) of f as:

$$f(x) \approx f(x_0) + C(x - x_0)$$

  where A is an $m \times n$ matrix.

- Thus, we are going to define the **total derivative** of f at x_0 as the matrix C such that:

$$\lim_{x \to x_0} \frac{\|f(x) - f(x_0) - C(x - x_0)\|}{\|x - x_0\|} = 0$$

Total Derivative - Definition

Definition 2.1

- Let $A \subset \mathbb{R}^n$ and $x \in int\,(A)$. A function $f : A \to \mathbb{R}^m$ is said to be **differentiable at** x iff $\exists$ an $m \times n$ real matrix C s.t.

$$\lim_{h \to 0} \frac{\|f(x+h) - f(x) - Ch\|}{\|h\|} = 0$$

  In this case, define the **(total) derivative of** f **at** x as the matrix C, denoted as $f'(x)$, or $Df(x)$.
- A function $f : A \to \mathbb{R}$ is said to be **differentiable** iff A is open and f is differentiable at any $x \in A$.
- Let $A_1 \subset int\,(A)$ be the set of points at which f is differentiable. Then the function $f' : A_1 \to \mathbb{R}^{mn}$ is called the **derivative (function)** of f.

Total Derivatives

- The derivative of a function from $\mathbb{R}^n \to \mathbb{R}^m$ is thus an $m \times n$ matrix C which we interpret as a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$, i.e., we could think of it as some mapping $\lambda(h)$.
- Intuitively, it is the matrix such that $f(x) + Ch$ approximates $f(x+h)$ well when $h \in \mathbb{R}^n$ is close to 0.
- For a function f from $\mathbb{R}^n$ to $\mathbb{R}$, the derivative reduces to a $1 \times n$ row vector called the **gradient** $(\nabla f(x))$

Some Properties

- The derivative is linear. If $f, g : \mathbb{R}^n \to \mathbb{R}^m$ and $\alpha \in \mathbb{R}$
  - $(f+g)' = f' + g'$
  - $(\alpha f)' = \alpha f'$
- If f differentiable at x, then f is continuous at x.

Introducing Partial Derivatives

- For a function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$, each coordinate $i \in \{1, \ldots, m\}$ of f can be regarded as a function f_i from A to $\mathbb{R}$.
- For instance, the function $f : \mathbb{R}^2 \to \mathbb{R}^2$

  $$f(x) = f(x_1, x_2) = (x_1 + x_2, x_1 - x_2)$$

  could be written as:

  $$f(x) = (f_1(x), f_2(x))$$

  where $f_1(x) = x_1 + x_2$ and $f_2(x) = x_1 - x_2$.
- f is differentiable at $x \in int\,(A)$ iff f_i is differentiable at x for each i, and furthermore we have

  $$f'(x) = \begin{bmatrix} \nabla f_1(x) \\ \nabla f_2(x) \\ \vdots \\ \nabla f_m(x) \end{bmatrix}$$

But first... An Example

- $f : \mathbb{R}^2 \to \mathbb{R}^2$, $f(x) = f(x_1, x_2) = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$. Is it differentiable at $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$?
- We have $f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}\right) = \begin{pmatrix} h_2 \\ h_1 \end{pmatrix}$
- We also have $f\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
- We are thus trying to find the matrix C such that

$$\lim_{h \to 0} \frac{\left\| \begin{pmatrix} h_2 \\ h_1 \end{pmatrix} - C \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\|}{\left\| \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} \right\|} = 0$$

- By setting $C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, the error term becomes 0 (no need to compute the limit). Our function is already linear, but the derivative exists!

Partial Derivatives

Definition 2.2

Let $A \subset \mathbb{R}^n$ and $x \in int\,(A)$. For a function $f : A \to \mathbb{R}^m$, its **partial derivative of the i-th coordinate w.r.t. the j-th argument at $x \in A$** is

$$\frac{\partial f_i}{\partial x_j}(x) := \lim_{h \to 0} \frac{f_i(x + he_j) - f_i(x)}{h}$$

if the right-hand side derivative exists.
The vector e_j above is the j-th canonical basis of $\mathbb{R}^n$, i.e. $e_j := (0, \ldots, 1, \ldots, 0)$.

Interpretation

- Basically, we are going back to the $\mathbb{R}$ to $\mathbb{R}$ case...
- The vector $x + he_j$ is a deviation from x only in the j-th argument. Therefore, intuitively, the partial derivative $\frac{\partial f_i}{\partial x_j}(x)$ measures the sensitivity of the i-th coordinate f_i of the multivariate function f wrt the j-th argument of x_j
- What are the partial derivatives of $f(x) = x_1^2 + x_1 x_2$?
- Partial derivatives provide a way to compute the total derivative of a function.

$f'(x)$ with Partial Derivatives

Theorem 2.3

Let $A \subset \mathbb{R}^n$ and $x \in int\,(A)$. If function $f : A \to \mathbb{R}^m$ is differentiable at x, then $\frac{\partial f_i}{\partial x_j}(x)$ exists for any $(i, j) \in \{1, \dots m\} \times \{1, \dots, n\}$, and furthermore we have

$$f'(x) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) & \cdots & \frac{\partial f_2}{\partial x_n}(x) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \frac{\partial f_m}{\partial x_2}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{bmatrix}$$

- Given this theorem, what is the total derivative (actually gradient) of $f(x) = x_1^2 + x_1 x_2$?

Technical Concerns

- We have seen that if f is differentiable, its partials exist and the Jacobian is just the matrix of partial derivatives.
- What if we only know that the partials exist? Is that enough for differentiability?
- Sadly, no... the function could behave nicely along the axes, but misbehave along other directions
- However, if the partials are also continuous, then the derivative exists (see Theorem 2.5). Almost every function we work with in economics will have continuous partial derivatives

Existence of Partial Derivatives Does Not Imply f Differentiable

- We can find a function f s.t. $\frac{\partial f_i}{\partial x_j}(x)$ exists for all (i,j) but f is not differentiable at x.
- Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$:

$$f(x,y) = \begin{cases} \frac{x^2 y}{x^4+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

- We apply definition of partial derivatives:

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = 0$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = 0$$

Existence of Partial Derivatives Does Not Imply f Differentiable

- Therefore, both partial exists at $(0, 0)$
- Nonetheless, the function is not differentiable at $(0, 0)$ as it is not continuous at $(0, 0)$.
- To see this, notice that f takes value $\frac{1}{2}$ along the path $y = x^2$, except for at the point $(0, 0)$, where it takes value 0.

Some Common Derivatives

- Being comfortable taking vector derivatives in one step can save you a lot of algebra (especially in econometrics). You should know these identities by heart :
- Let $f : \mathbb{R}^n \to \mathbb{R}^m$ with $f(x) = Ax$ where A is an $m \times n$ matrix:

$$f'(x) = A$$

- Let $f : \mathbb{R}^n \to \mathbb{R}$ with $f(x) = x'Ax$ where A is an $n \times n$ matrix:

$$f'(x) = x'(A + A')$$

  If A is symmetric, $f'(x) = 2x'A$
- If $f, g : \mathbb{R}^n \to \mathbb{R}$ and $h(x) = f(x)g(x)$:

$$h'(x) = f'(x)g(x) + g'(x)f(x)$$

Chain Rule

- Let's state the chain rule for a single variable function:

Chain Rule

Let S be a subset of $\mathbb{R}$, and $f : S \to \mathbb{R}$. Let T be a set s.t. $f(S) \subset T \subset \mathbb{R}$, and $g : T \to \mathbb{R}$. If f is differentiable at x, and g is differentiable at $f(x)$, and we have

$$(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$$

- For instance, suppose we want to take the derivative of $h(x) = \ln(x^2 + 1) = g(f(x))$, where $f(x) = x^2 + 1$ and $g(y) = \ln y$.
- Then $f'(x) = 2x$ and $g'(y) = \frac{1}{y}$.
- Thus, $h'(x) = g'(x^2 + 1) \cdot 2x = \frac{2x}{x^2+1}$

Chain Rule in multivariate Functions

- We can extend the chain rule for multivariate functions.

Chain Rule

Let $S \in \mathbb{R}^n$, $x \in int(S)$, and $f : S \to \mathbb{R}^m$. Let T be s.t. $f(S) \subset T \subset \mathbb{R}^m$ and $f(x) \in int(T)$, and let $g : T \to \mathbb{R}^k$. If f is differentiable at x, and g is differentiable at $f(x)$, then $g \circ f : S \to \mathbb{R}^k$ is differentiable at x. Furthermore, we have

$$(g \circ f)'(x) = g'(f(x)) \cdot f'(x)$$

- In the equation above, the $\cdot$ on the right-hand side is the matrix multiplication. Because $g'(f(x))$ is an $k \times m$ matrix, and $f'(x)$ is an $m \times n$ matrix, their product $g'(f(x)) \cdot f'(x)$ is a $k \times n$ matrix, which is exactly the size $(g \circ f)'(x)$ should have.

Chain Rule - Example

- Consider the following example:
- Utility u depends on consumption c and hours worked h
- However, c and h depend on the going wage w . Define $x(w) : \mathbb{R} \to \mathbb{R}^2$ by $x(w) = (c(w), h(w))$. This function assigns consumption and hours worked for any level of wage.
- We thus define our utility as $v(w) = u(x(w))$. Chain rule says $v'(w) = u'(x(w)) \cdot x'(w)$:

$$
\begin{aligned}
v' &= \begin{pmatrix} \frac{\partial u}{\partial c} & \frac{\partial u}{\partial h} \end{pmatrix} \begin{pmatrix} \frac{\partial c}{\partial w} \\ \frac{\partial h}{\partial w} \end{pmatrix} \\
&= \frac{\partial u}{\partial c}\frac{\partial c}{\partial w} + \frac{\partial u}{\partial h}\frac{\partial h}{\partial w}
\end{aligned}
$$

Higher Derivatives: Single Variable

- For a function $f : \mathbb{R} \to \mathbb{R}$, the **second derivative** of f at x is the derivative of f' at x:

$$f''(x) = \lim_{h \to 0} \frac{f'(x+h) - f'(x)}{h}$$

- The second derivative measures the change in the slope per unit change in x:
  - If $f''(x) > 0$ then the derivative is (locally) increasing in x
  - If $f''(x) < 0$ then the derivative is (locally) decreasing in x

Higher Derivatives: $\mathbb{R}^n$

- Recall that for a function $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ we know that its gradient at $x \in int(A)$ is equal to the vector of partial derivatives: i.e.

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \ldots, \frac{\partial f}{\partial x_n}(x) \right)$$

Higher Derivatives: $\mathbb{R}^n$

- The second derivative of the real-valued function f at x is also known as the **Hessian matrix** of f at x, denoted as $H_f(x)$:

$$H_f(x) := f''(x) = (\nabla f)'(x) = \begin{bmatrix} \left(\nabla \frac{\partial f}{\partial x_1}\right)(x) \\ \left(\nabla \frac{\partial f}{\partial x_2}\right)(x) \\ \vdots \\ \left(\nabla \frac{\partial f}{\partial x_n}\right)(x) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{\partial\left(\frac{\partial f}{\partial x_1}\right)}{\partial x_1}(x) & \frac{\partial\left(\frac{\partial f}{\partial x_1}\right)}{\partial x_2}(x) & \cdots & \frac{\partial\left(\frac{\partial f}{\partial x_1}\right)}{\partial x_n}(x) \\ \frac{\partial\left(\frac{\partial f}{\partial x_2}\right)}{\partial x_1}(x) & \frac{\partial\left(\frac{\partial f}{\partial x_2}\right)}{\partial x_2}(x) & \cdots & \frac{\partial\left(\frac{\partial f}{\partial x_2}\right)}{\partial x_n}(x) \\ \vdots & \vdots & & \vdots \\ \frac{\partial\left(\frac{\partial f}{\partial x_n}\right)}{\partial x_1}(x) & \frac{\partial\left(\frac{\partial f}{\partial x_n}\right)}{\partial x_2}(x) & \cdots & \frac{\partial\left(\frac{\partial f}{\partial x_n}\right)}{\partial x_n}(x) \end{bmatrix}$$

Order of differentiation matters?

- The cross partial

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x) := \frac{\partial\left(\frac{\partial f}{\partial x_i}\right)}{\partial x_j}(x)$$

and the cross partial

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) := \frac{\partial\left(\frac{\partial f}{\partial x_j}\right)}{\partial x_i}(x)$$

are conceptually very different when $i \neq j$. However, they are equal if f is twice continuously differentiable at x, and this result is usually known as Young's theorem or Schwarz's theorem.

Young; Schwarz

Young-Schwarz Theorem

Let $A \subset \mathbb{R}^n$ and $x \in int\,(A)$. If function $f : A \to \mathbb{R}$ is C^2 at x, then for any $i, j \in \{1, \ldots, n\}$ both $\frac{\partial^2 f}{\partial x_j \partial x_i}(x)$ and $\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$ exists and

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(x) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$$

Taylor Series - "Intuitively"

- Suppose you want to approximate $f : \mathbb{R} \to \mathbb{R}$ by a polynomial around x_0:

$$h(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_n(x - x_0)^n$$

- Two intuitive criteria for a "good" approximation are:
  1. $h(x_0) = f(x_0) \implies a_0 = f(x_0)$
  2. The first n derivatives of h should match those of f at x_0
- Differentiating repeatedly gives $h^k(x_0) = k!a_k$. Thus the Taylor series expansion of order n of f around x_0 is:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \cdots + \frac{1}{n!}f^n(x_0)(x - x_0)^n$$

We are going to be more precise next class...