Math Review
Summer 2017

Topic 3

3. Calculus, differentiation and integrals

3.1. One-variable calculus and rules of differentiation

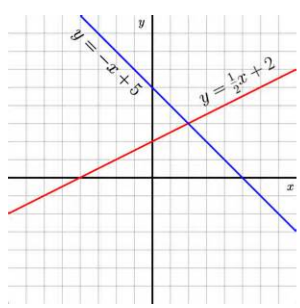
Calculus is a key tool for the micro series. We review the key definitions and work through a bunch of examples and exercises diligently. One gets a lot more by doing rather than reading through solutions. As you progress through your micro courses, you will likely forget some of the rules here and there, don't feel discouraged, just look at some reference material and you will be fine.

3.1.1. Function

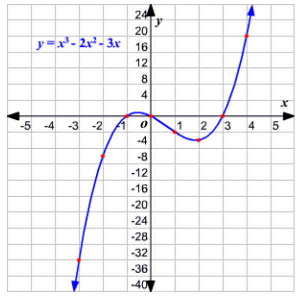
Let's define a function in a one-variable calculus context.

Function: A function, $f: A \rightarrow B$ takes $x \in A$ and produces $f(x) \in B$. A function may assign multiple number to one number, i.e., $f : \mathbb{R}^n \rightarrow \mathbb{R}$, but in *one-variable calculus* a function simply maps one number to another number, i.e., $f: \mathbb{R} \rightarrow \mathbb{R}$. If f maps distinct points to distinct values, then f is *one-to-one* (or injective). Some examples of functions are:

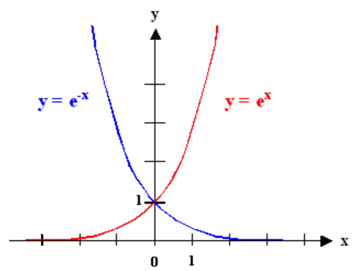
Linear function: $f(x) = y = ax + b$, where a is the slope and b is the intercept



Polynomial function: $f(x) = y = a_0 + a_1 x + \cdot + a_{k-1}x^{k-1} + a_k x^k$

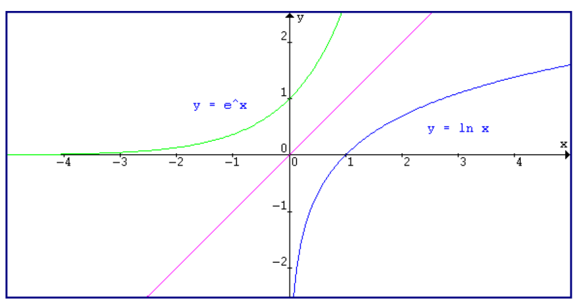


Exponential function: $f(x) = y = ae^{bx}$



Logarithmic function: $f(x) = y = \log_b(x)$

Key one is that of the $ln(x)$:



There are other types of functions, but they are rarely used (at least not in your first year study) – for example trigonometric functions of the form $a\, sinx + b\cos x$.

Here y would be the *dependent variable*, also referred to as the *endogenous* variable and x the *independent* or *exogenous* variable.

3.1.2. Derivatives

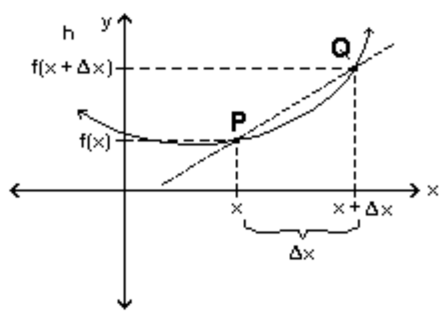
You will often hear about rate of change or the marginal effect. For example, when we have a production function for a good, $y = f(x)$, we are usually interested in how a change in the inputs, $\Delta\, x$, will affect the production of the good $\Delta\, f\ = \Delta\, y$. The *derivative* of a function contains this information.

We will not go over too many details on derivatives as a function of limit, but it may helpful to recall that a derivative can be expressed as a limit.

The derivative of a function $f\colon \mathbb{R} \to \mathbb{R}$ at $x = \hat{x}$ is given by:

$$f' = \frac{df}{dx}(\hat{x}) = \lim_{h_n \to 0} \frac{f(\hat{x} + h_n) - f(\hat{x})}{h_n}$$

You will be able to take derivatives directly given the rules that you are familiar with (and that we will review today) but it is helpful to see an example of what is at work in the background. Taking derivatives are so easy that we rarely think back to where the basic formulas and rules originated.



Example. Use the limit definition to compute the derivative, $f'(x)$, for:

$$f(x) = \frac{1}{2}x - \frac{3}{5}$$

$$f' = \frac{df}{dx}(\hat{x}) = \lim_{\Delta x \to 0} \frac{f(\hat{x} + \Delta x) - f(\hat{x})}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{\left(\frac{1}{2}(x + \Delta x) - \frac{3}{5}\right) - \left(\frac{1}{2}x - \frac{3}{5}\right)}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{\left(\frac{1}{2}x + \frac{1}{2}\Delta x - \frac{3}{5}\right) - \frac{1}{2}x - \frac{3}{5}}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{\frac{1}{2}\Delta x}{\Delta x}$$

$$= \frac{1}{2}$$

$\mathcal{Q}$: Can you try with $f(x) = x^2$

$\mathcal{A}$:

$$f' = \frac{df}{dx}(\hat{x}) = \lim_{\Delta x \to 0} \frac{f(\hat{x} + \Delta x) - f(\hat{x})}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{(x + \Delta x)^2 - x^2}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{x^2 + 2x\Delta x + (\Delta x)^2 - x^2}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{2x\Delta x + (\Delta x)^2}{\Delta x}$$

$$= \lim_{\Delta x \to 0} 2x + \Delta x$$

$$= 2x$$

The function f is differentiable at $\hat{x}$ if and only if the limit exists. The function f is differentiable if and only if it is differentiable at every point $x \in \mathbb{R}$.

Let's quickly recall our definition of continuous functions from *§ 2.7:*

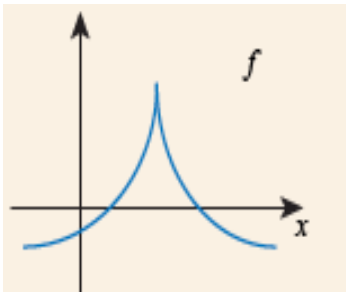
<u>Continuous function</u>: A function $f: \mathbb{R} \to \mathbb{R}$ is continuous at a point at $p \in \mathbb{R}$ if and only if, for every $\varepsilon > 0$, there exists $\delta > 0$, such that:

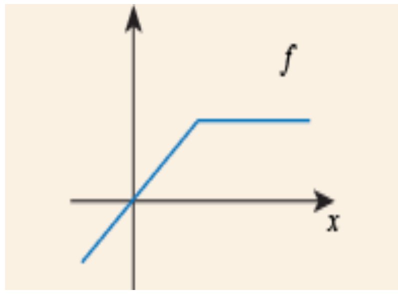
$$|x - p| < \delta \text{ } implies \text{ } |f(x) - f(p)| < \epsilon$$

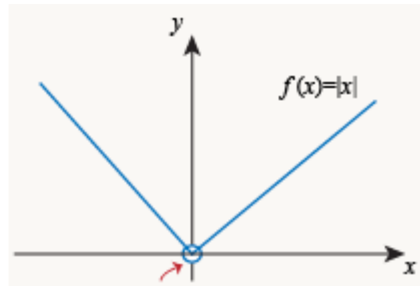
If f is differentiable at x then f is continuous at x. So the differentiability of f at x is a <u>sufficient</u> condition to show that f is continuous at x. By the <u>contrapositive</u> we also have: If f is not continuous at x then f is not differentiable at x.

Q: Can a function be continuous at x but not differentiable at x?

A: Yes, kinks and sharp ends.







Thus, we can say that a function f is differentiable if it is both continuous and "smooth".

The derivative $f'(x)$ is the rate of change in $f(x)$ and can be expressed as:

$$\frac{\partial f(x)}{\partial x} = f'(x)$$

- If $f'(x) > 0$ for $x \in [a, b]$, then $f(\cdot)$ is increasing on $[a, b]$.
- If $f'(x) < 0$ for $x \in [a, b]$, then $f(\cdot)$ is decreasing on $[a, b]$.
- If $f'(x) \geq 0$ for all $x \in \mathbb{R}$, then we say that $f(\cdot)$ is *monotonically increasing* (or non-decreasing). If $f'(x) > 0$ for all $x \in \mathbb{R}$, then $f(\cdot)$ is *strictly increasing*.
- The point $\hat{x} \in \mathbb{R}$ is a critical point of the continuous and differentiable function $f(x)$ if $f'(x) = 0$ or $f'(x) = \emptyset$.

3.1.2 Second derivative

If the (first) derivative is a differentiable function, we can take its derivative, and get the *second derivative* which is expressed as:

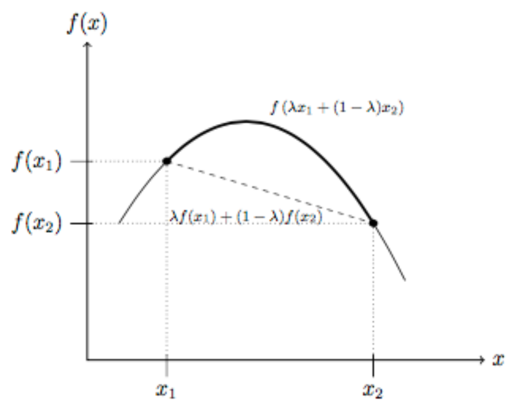
$$\frac{\partial^2 f(x)}{\partial x^2} = f''(x)$$

- All higher order derivatives are defined in the same way. We also say that a function is *continuously differentiable* if it is *continuous*, *differentiable*, and the <u>derivative</u> is a *continuous function.*

- If a function possesses continuous derivatives $f'(x), f''(x), \ldots f^n(x)$ it is called n-times continuosly differentiable or C^n.

3.1.3. Derivatives, concavity and convexity of functions

Since we are on second derivatives, these are some useful results that you will encounter quite a bit in Micro.

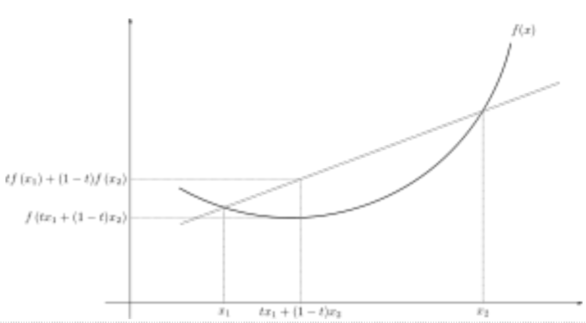


The function $f: \mathbb{R} \to \mathbb{R}$ is concave if and only if $\boldsymbol{f''(x) \leq 0}$ for all x. As you can see in the graph, if for any two points $x_1, x_2 \in \mathbb{R}$ and any $\lambda \in [0,1]$ we have

$f(\lambda x_1 + (1-\lambda)x_2) \geq \lambda f(x_1) + (1-\lambda)f(x_2)$, then $f(x)$ is concave.

$\mathcal{Q}$: How would you define a convex function? Draw a diagram to follow the logic.

$\mathcal{A}$: The function $f: \mathbb{R} \to \mathbb{R}$ is convex if and only if $\boldsymbol{f''(x) \geq 0}$ for all x. As in the graph, if for any two points $x_1, x_2 \in \mathbb{R}$ and any $t \in [0,1]$ we have
$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$, then $f(x)$ is convex.



If $f''(x) > 0$ for all x, then $f(\cdot)$ is strictly convex. If $f''(x) < 0$ for all x, then $f(\cdot)$ is strictly concave.

3.1.4 Rules of differentiation and practice

Some basic rules of differentiation are as follows:

1. For constant α: $\frac{\partial \alpha}{\partial x} = 0$
2. For sums: $\frac{\partial}{\partial x}\,[f(x) \pm g(x)] = f'(x) \pm g'(x)$
3. Power rule: $\frac{\partial}{\partial x}\,(\alpha x^n) = n\alpha x^{n-1}$
4. Product rule: $\frac{\partial}{\partial x}\,[f(x)g(x)] = f(x)g'(x) + f'(x)g(x)$
5. Quotient rule: $\frac{\partial}{\partial x}\left[\frac{f(x)}{g(x)}\right] = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$
6. Chain rule: $\frac{\partial}{\partial x}\,[f\,(g(x))] = f'(g(x))g'(x)$

Let's practice finding the derivatives of the following:

$a. (x^2 - 5)(x^3 - 2x + 3)$

$$Ans: 5x^4 - 21x^2 + 6x^2 + 10$$

$b. \frac{x^2 + 1}{5x - 3}$

$$Ans: \frac{5x^2 - 6x - 5}{(5x - 3)^2}$$

$c. (x^3 + 4)^4$

$$Ans: 12x^2(x^3 + 4)^3$$

$d. \sqrt{x^3 + 2x + 1}$

$$Ans: \frac{3x^2 + 2}{2\sqrt{x^3 + 2x + 1}}$$

$e. \left(\frac{x + 6}{x + 5}\right)^{1/4}$

$$Ans: -\frac{1}{4}\left(\frac{x + 6}{x + 5}\right)^{-\frac{3}{4}} . \frac{1}{(x + 5)^2}$$

Professor Glewwe and Coggins have many instances where they extensively use derivatives with logs. Remember that if $y = \ln u, then \frac{\partial y}{\partial x} = \frac{u'}{u}$ or if $h = \ln\ g(x), then \frac{\partial h}{\partial x} = \frac{g(x)'}{g(x)}$

Find the derivative of the following:

$f. 2\ln{(3x^2 - 1)}$

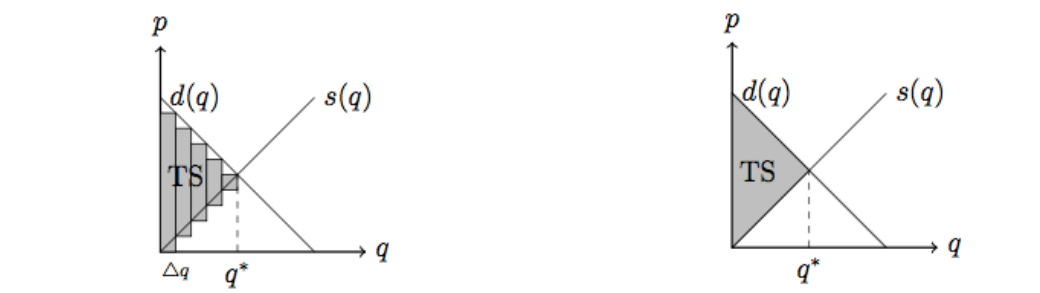
$$Ans: \frac{12x}{3x^2 - 1}$$

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

$$Ans: \frac{x^2 - 2x - 1}{(x^2 + 1)(x - 1)}$$

3.2 Integrals and related properties

In economics we are often interested in the integral of a function. Recall back to your basic demand and supply curves. The area under the curves gives us the total surplus. This area underneath the curves can be obtained using the actual integral or sometimes using approximations (approximate an integral by adding up the area of rectangles). You will see both in micro.



(a) Approximation: $TS \approx \sum (d(q) - s(q)) \triangle q$ (b) Actual: $TS = \int_0^{q^*} d(q) - s(q)dq$

If we are *approximating* using the area of the rectangles, we are really finding the integral through the following:

$$\int_a^b f(x)dx = \sum f(x) \triangle x$$

The *actual* integral of a function is the limit of the above function as $\triangle x$ approaches zero. We have:

<u>Integral</u>. The integral of a function $f: \mathbb{R} \rightarrow \mathbb{R}$ on $[a, b]$ is given by:

$$\int_a^b f(x)dx = \lim_{x \to \Delta x} \sum f(x) \triangle x$$

The function f is integrable on $[a, b]$ if and only if the limit exists. We can also call the integral the *antiderivative* of f. If f is continuous on $[a, b]$ then f is also integrable on $[a, b]$.

Properties:

Suppose that f and g are integrable functions. Let $a, c \in \mathbb{R}$ be arbitrary constants. Then:

(i) $\int af(x)dx = a \int f(x)dx$

(ii) $\int (f(x) + g(x))\, dx = \int f(x)dx + \int g(x)dx$

The Fundamental Theorem of Calculus Theorem. Let f be a continuous function on the open interval $[a, b]$. If $f(x) = F'(x)$, then:

$$\int_a^b f(x)dx = F(b) - F(a).$$

We also have the following three properties:

(iii) $\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|dx$

(iv) $\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx$

(v) If $f(x) \leq g(x)$ and $a < b$, then $\int_a^b f(x)dx \leq \int_a^b g(x)dx$

Remember that if the integration does not have a specified range, then we need to include a constant of integration to the solution:

$\int f(x)dx = F(x) + C$, where C is the constant of integration.

Remember that $\int e^x dx = e^x + C$.

Q: What is the $\int \frac{1}{x} dx = ?$

A: $\int \frac{1}{x} dx = lnx + C$

Let's work through a simple example to refresh our memory of integrals.

Example. Calculate the integral of $\int (4x^2 \ + \ x^1 \ - \frac{3}{x})\, dx$ where $F\,(1) \ = \ 0$.

$$\int (4x^2 \ + \ x^{\frac{1}{2}} \ - \frac{3}{x})\, dx = \ 4\int x^2 dx + \int x^{\frac{1}{2}}\, dx - 3\int \frac{1}{x}$$
$$= \frac{4}{3}x^3 + \frac{2}{3}x^{\frac{3}{2}} + 3\; lnx + C$$

Given that $F\,(1) \ = \ 0$, we have $\frac{4}{3} + \frac{2}{3} + C = 0; C = -2$

We can write $F\,(x) = \frac{4}{3}x^3 + \frac{2}{3}x^{\frac{3}{2}} + 3\; lnx - 2$

3.2.1. Integration by parts

Integration by parts is used a little less often but you may see it sometimes. It corresponds to the Product Rule for differentiation.

$$\int f\, g' dx = fg - \int gf' dx$$

Others may remember it as letting $u = f$ and $v = g$, then by substitution, we have:

$$\int u\, dv = \ u\, v - \int v\, du$$

Example. Find $\int \ln x\, dx$. Some may know this final result, but let's try it using integration by parts.

There is not much you can pick for u and dv, so this is pretty straightforward.

Let $u = \ln x, \ du = \frac{1}{x}$
$dv = dx, v = x$

Using integration by parts:

$$\int u\, dv = \ u\, v - \int v\, du$$
$$= \mathrm{x} \ln x - \int x \frac{dx}{x}$$

$$x\ lnx - x + C$$

Q: Let's have you try a slightly more involved example. Find: $\int x^2 e^x$. *Hint, recall sometimes you have to repeat an integration by parts.*

$\mathcal{A}$:

$$\int udv = uv - \int vdu$$

Let $u = x^2; du = 2x\ dx$
Let $dv = e^x; v = e^x$

$$= x^2e^x - \int 2xe^x dx$$
$$= x^2e^x - 2\int xe^x dx$$

Let $u = x; du = dx$
Let $dv = e^x; v = e^x$

$$= x^2e^x - 2[xe^x - \int e^x dx\ ]$$
$$= x^2e^x - 2xe^x + e^x + C$$

3.2.2. Leibniz's Rule

Finally, it might be helpful to reference the Leibniz's Rule of differentiating integrals. It is a little complex looking, but an example may help.

Leibniz's Rule:

$$\frac{d}{dt}\int_{a(t)}^{b(t)} f(x,t)dx = \frac{db(t)}{dt}f(b(t),t) - \frac{da(t)}{dt}f(a(t),t) + \int_{a(t)}^{b(t)} \frac{\partial f(x,t)}{\partial t}dx$$

Example. Calculate $\frac{d}{dy}\int_{2+y}^{y^2}(x+y)^2 dx$

Let's think of the equivalence from the Rule,

$$t \to y$$
$$a(t) \to 2 + y$$

$$b(t) \to y^2$$
$$f(x,t) \to (x+y)^2$$
$$\frac{\partial f(x,t)}{\partial t} \to \frac{\partial f(x,t)}{\partial y} = 2(x+y)$$

$$\frac{d}{dy}\int_{2+y}^{y^2}(x+y)^2 dx = \frac{dy^2}{y}(y^2+y) - \frac{d(2+y)}{dy}\left((2+y)+y\right)^2 + \int_{2+y}^{y^2} 2(x+y)dx$$
$$= 2y\,(y^2+y) - (2+2y)^2 + (x^2+2xy\Big|_{2+y}^{y^2})$$
$$= 2y\,(y^2+y) - (2+2y)^2 + y^4 + 2y^3 + (2+y)^2 + 2y(2+y)$$

You can further evaluate this, but you get the point!

$\mathcal{Q}$: What happens to the expression for Leibniz Rule when $a(t)$ and $b(t)$ are scalars a and b instead?

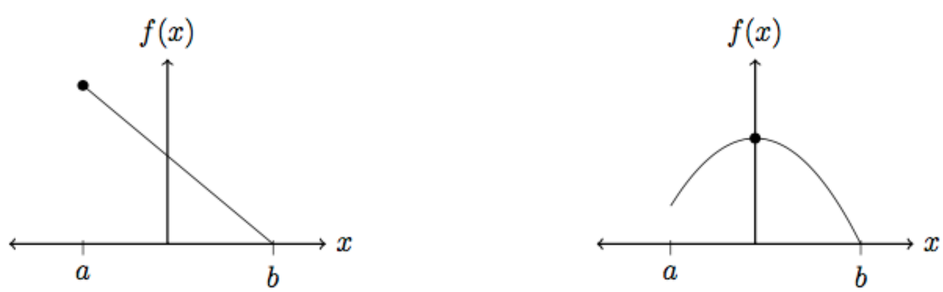
$\mathcal{A}$: Only the last term remains, actually this is how you will use if often.

$$\frac{d}{dx}\int_a^b f(x,y)dy = \int_a^b \frac{\partial f(x,y)}{\partial x} dy$$

3.3. Maxima and minima

3.3.1. Local maxima

- The point $x*$ is a local maximum on the interval $[a,b]$ if $f(x*) \geq f(x)$ for all $x \in [a,b]$.
- If $x*$ is an interior max of $f(\cdot)$, then $x*$ is a *critical point* of $f(\cdot)$.



(a) Boundary max (b) Interior max

At a critical point $x*$ on the domain $[a, b]$, the second derivative can be used to check whether it is a maximum or a minimum.

1. $f'(x*) = 0; f''(x) < 0 \rightarrow$ Maximum
2. $f'(x*) = 0; f''(x) > 0 \rightarrow$ Minimum
3. $f'(x*) = 0; f''(x) = 0 \rightarrow$ Indeterminate

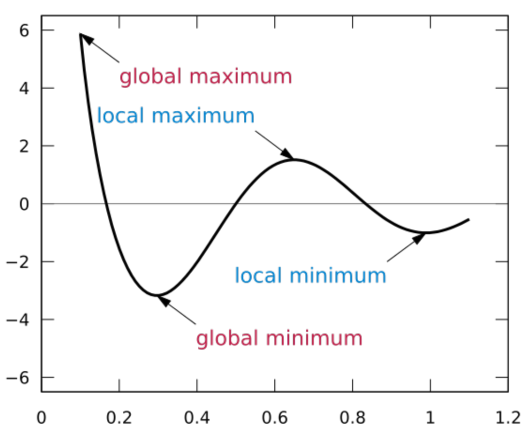
How do you remember these relations? This is how I do! (show)

3.3.2. Global maxima

The point $x*$ is a global maximum if $f(x*) \geq f(x)$ for all x in the doman of $f(\cdot)$. Note that a global maximum need **not** be a *critical point*: The function $f(\cdot)$ has a global maximum when:

Theorem. Let $f(x)$ be a twice differentiable function on the domain $\mathbb{R}$. If $x*$ is a local maximum of $f(\cdot)$ and $x*$ is the only critical point of f on A, then $x*$ is a global maximum of f on A.

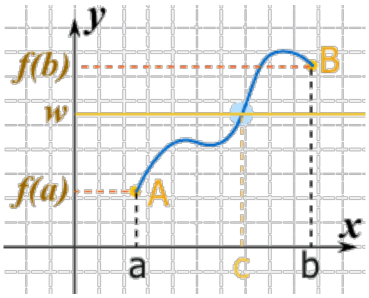
Let's try an example. Which points are the local maxima and minima? (If they exist).



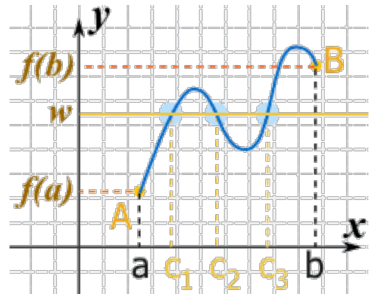
3.3.3 Other useful theorems:

Intermediate Value Theorem. Let $f(\cdot)$ be a continuous function on the closed interval $[a, b]$. And let N be any given number between between $f(a)$ and $f(b)$, where. If $f(a) \neq f(b)$. Then there exists a number c in (a, b), such that $f(c) = N$

Explanation: This is getting at the essence of not lifting your pencil when drawing a graph. w= N



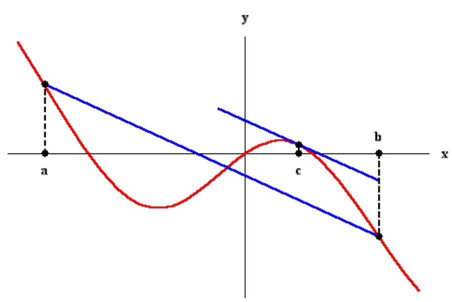
Can be more than 1 c:



Mean Value Theorem: Suppose that $f(\cdot)$ is continuous on $[a,b]$ and differentiable on (a,b). Then for some $c \in (a,b)$,

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

Explanation: This is really the slope of the line passing through $(a, f(a))$ *and* $(b, f(b))$. So the conclusion of the Mean Value Theorem states that there exists a point $c \in (a.b)$such that the tangent line is parallel to the line passing through $(a, f(a))$ *and* $(b, f(b))$.



Weierstrass' Theorem. A continuous function, $f(\cdot)$, on the closed and bounded interval $[a,b]$ attains both a local maximum and minimum.

Taylor's Theorem. The Taylor series if the function f at a is given by:

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \frac{1}{3!}f'''(a)(x-a)^3 + \cdots$$
$$= \sum_{k=0}^{n} \frac{f^k(a)}{k!}(x-a)^k$$

Often, when you need to use Taylor series approximation, you will work with a manageable order of the series. For example, zeroth, the first, second, and third order of the Taylor polynomial are:

$$P_0(x) = f(a)$$
$$P_1(x) = f(a) + f'(a)(x - a),$$
$$P_2(x) = f(a) + f'(a)(x - a) + \frac{f''(a)(x - a)^2}{2!}$$

$$P_3(x) = f(a) + f'(a)(x - a) + \frac{f''(a)(x - a)^2}{2!} + \frac{f'''(a)(x - a)^3}{3!}$$

Let's think about what is happening here. You are approximating $f(x)$ and as you increase your order, you get closer and closer.

For example, $P_1(x)$ is the same as the linear approximation of $f(x)$ centered at $x = a$, so it is often called "the first-order approximation of $f(x)$ at (or near) $x = a$." $P_2(x)$ is then called the quadratic, or second-order approximation, $P_3(x)$ the cubic, or third-order approximation,
and so on.

Example. Find $P_0(x), P_2(x), \cdot \cdot \cdot, P_5(x)$ at x = 0 for the function $f(x) = e^x$.

$$f(x) = f'(x) = f''(x) = f'''(x) = f^{(4)}(x) = f^{(5)}(x) = e^x$$

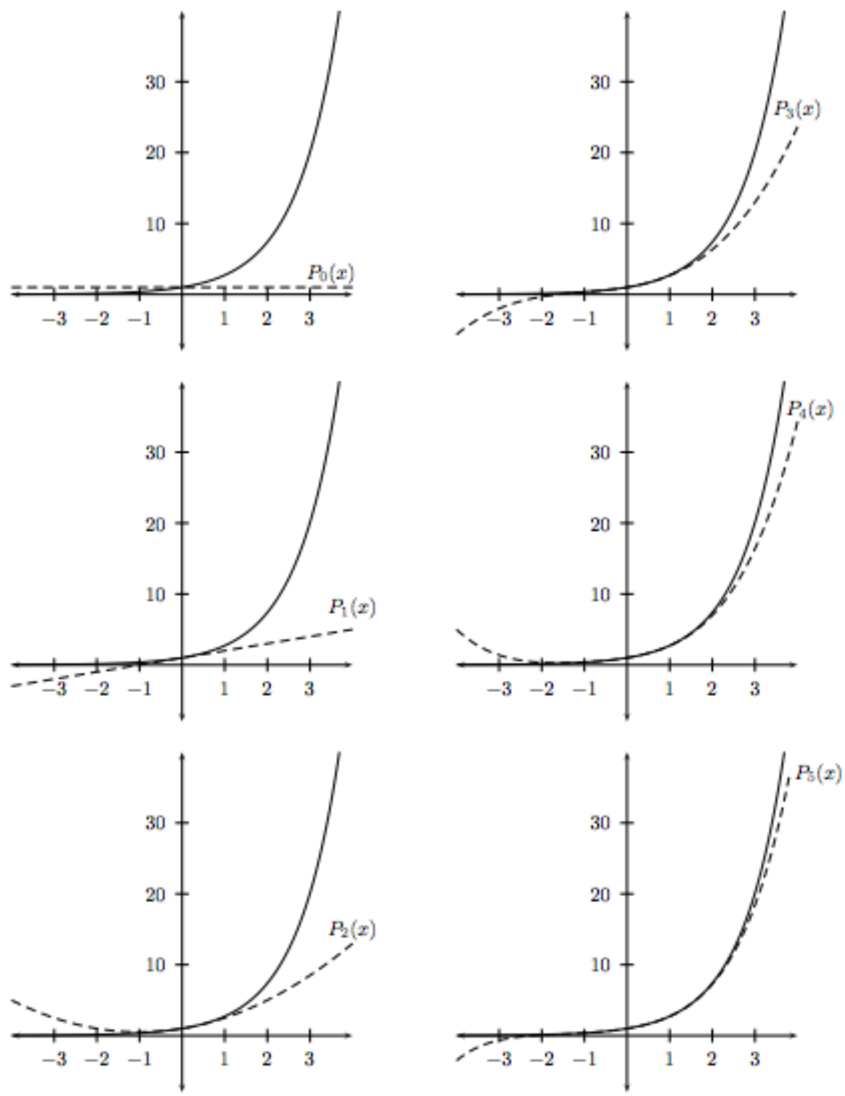
Let's center this at $a = 0$. From the definition:

$$\begin{aligned} P_5(x) = {} & f(0) + f'(0)(x - 0) + \frac{f''(0)(x - 0)^2}{2!} + \frac{f'''(0)(x - 0)^3}{3!} \\ & + \frac{f'^{(4)}(0)(x - 0)^4}{4!} + \frac{f'^{(5)}(0)(x - 0)^5}{5!} \\ & = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \frac{1}{5!}x^5 \end{aligned}$$

You can easily identify that:

$$P_0(x) = 1\ ;\ P_1(x) = 1 + x; \ldots;\ P_5(x) = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \frac{1}{5!}x^5$$

So graphically, how good are we getting with the higher order?



Note that we used a nice example, but things are not always as neat. With varying functional forms and values of a, the approximation may not be as good as it sounds. Depending on what field courses you take you may or may not see a larger expose on Taylor series expansions. For now, this should be quite plenty.

3.4. Review of Euclidean n-space

A lot of what we do is in the n-space. If we are given $x \in \mathbb{R}^n$, it can be viewed as a vector:

$$\boldsymbol{x} = (x_1, x_2, \dots, x_n) \text{ or } \boldsymbol{x}' = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

All your addition and subtraction operations need to be undertaken element by element. The *dot product* or *inner dorm* is given by:

$$\boldsymbol{x} \cdot \boldsymbol{y} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

3.5. Multi-variable calculus

Now we move to function of several variables.

You will work with utility functions where individuals get utility from multiple commodities, $x_1, x_2, \dots, x_n$, with prices $p_1, p_2, \dots, p_n$. Or in production, you may have a production function where a particular output is made from several inputs. Consider a straightforward Cobb-Douglas production function using inputs x_1 and x_2: $f(x_1, x_2) = kx_1^a x_2^b$

A function from $\mathbb{R}^n$ to $\mathbb{R}$ takes the numbers $(x_1, x_2, \dots, x_n)$ in the domain of the function and assigns a number $y \in \mathbb{R}$, i.e. $f(x_1, x_2, \dots, x_n) = y$. Note that for each x_i where $i = (1,2, \dots, n)$ we have $x_i \in \mathbb{R}$.

A function from $\mathbb{R}^n$ to $\mathbb{R}^m$ takes the numbers $(x_1, x_2, \dots, x_n)$ in the domain of the function and assigns a number $(y_1, y_2, \dots, y_m)$, where $y_i \in \mathbb{R}$, for all $i = (1,2, \dots, m)$, i.e. $f(x_1, x_2, \dots, x_n) = (f_1(x_1,, x_2, \dots, x_n), f_2\big(x_1,, x_2, \dots, x_n, \dots, f_n(x_1,, x_2, \dots, x_n)\big) = (y_1, y_2, \dots, y_m)$.

Let's consider a useful mapping that you will see often in Micro. You have m consumers, k commodities, and a utility mapping which is a function from $\mathbb{R}^{km}$ to $\mathbb{R}^m$.

$$u\big((x_1^1, x_2^1, \dots, x_k^1), \dots, (x_1^m, x_2^m, \dots, x_k^m)\big) = (u^1(x_1^1, x_2^1, \dots, x_k^1), \dots, u^m(x_1^m, x_2^m, \dots, x_k^m))$$

3.6. Partial, total and higher order derivatives

3.6.1. Partial Derivatives

Partial Derivatives: A function on n variables can be thought to have n partial slopes, each giving only the rate at which y would change if one x_i, alone, where to change.

Let $y = f\,(x_1, x_2, \ldots x_n)$. The partial derivatives of f with respect to x_i is defined as:

$$\frac{\partial f}{\partial x_i} = \lim_{h \to 0} \frac{f(x_1 + x_2 + \cdots + x_i + h, \ldots, x_n) - f(x_1 + x_2 + \cdots + x_n)}{h}$$

Here is an example:

Example. Given $f(x_1, x_2) = x_1^2 + 3x_1x_2 - x_2^2$. Compute the partial derivative with respect to each element:

(a) $\frac{\partial f(x_{1,} x_2)}{\partial x_1}$

$$\frac{\partial f\left(x_{1,} x_2\right)}{\partial x_1} = 2x_1 + 3x_2$$

(b) $\frac{\partial f(x_{1,} x_2)}{\partial x_1}$

$$\frac{\partial f\left(x_{1,} x_2\right)}{\partial x_2} = 3x_1 - 2x_2$$

Taking partial derivative is not particularly harder than regular derivatives you have been taking. For example, if you are taking the derivative with respect to x_1, you can treat x_2 as a constant.

3.6.2. Total Derivatives

The total derivative of a function $f : \mathbb{R}^n \to \mathbb{R}$ tells us how f changes as we allow x_1 through x_n to change *simultaneously*. The total differential f at $\hat{\boldsymbol{x}}$ can be approximated as follows:

$$df = \frac{\partial f}{\partial x_1} * (\hat{\boldsymbol{x}}) dx_1 + \cdots + \frac{\partial f}{\partial x_n} * (\hat{\boldsymbol{x}}) dx_n$$

You may also encounter Df which is often called the *Jacobian derivative* of f at $\hat{\boldsymbol{x}}$. The Jacobian of $f : \mathbb{R}^{\boldsymbol{n}} \to \mathbb{R}^{\boldsymbol{m}}$ is:

$$J = Df = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

You may also see total derivatives used in the context of gradients. The gradient of f at $\hat{x}$ is the transpose of the derivative of f itself:

$$\nabla f(\hat{x}) = grad\ f(\hat{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\hat{x}) \\ \frac{\partial f}{\partial x_2}(\hat{x}) \\ \vdots \\ \frac{\partial f}{\partial x_n}(\hat{x}) \end{bmatrix}$$

It can be hard to picture the difference between some of these concepts. Mainly, the Jacobian and gradient vectors can be confusing. An example can help:

Suppose, $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x_1, x_2) = x^2 + y$, then $grad\ f(\hat{x})$ is a $2\ x\ 1$ matrix given as:

$$grad\ f(\hat{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\hat{x} \\ \frac{\partial f}{\partial x_2}(\hat{x}) \end{bmatrix}$$

$$grad\ f(\hat{x}) = \begin{bmatrix} 2x \\ 1 \end{bmatrix}$$

Now suppose, $f: \mathbb{R}^2 \to \mathbb{R}^2$, $f(x_1, x_2) = (x^2 + y, y^3)$, then the $Jacobian$ is the $2\ x\ 2$ matrix:

$$J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial y} \end{bmatrix}$$

$$J = \begin{bmatrix} 2x & 1 \\ 0 & 3y^2 \end{bmatrix}$$

Total derivative with chain rule:

Often, you will encounter total derivative in combination with the chain rule. Let's go through simple examples of both and then I will pull an example from the Production mini.

Example. Let $w = x^3yz + xy + z + 3$

Find the total derivative.

$$dw = \frac{\partial w}{\partial x}dx + \frac{\partial w}{\partial y}dy + \frac{\partial w}{\partial z}dz$$

$$= (3x^2yz + y)dx + (x^3z + x)dy + (x^3y + 1)dz$$

Now, you are given that $x = 2t + 1, y = 4t^2$, and $z = \ln t$

Find $\frac{dw}{dt}$

You can substitute all these $t's$ and get your answer, but as in Micro you will be dealing with functions that are often unknown, let's just denote that $x = x(t), y = y(t), z = z(t)$.

Then: $\frac{dw}{dt} = \frac{\partial w}{\partial x}\frac{dx}{dt} + \frac{\partial w}{\partial y}\frac{dy}{dt} + \frac{\partial w}{\partial z}\frac{dz}{dt}$. All these individual elements are easier to compute.

$$= [(3x^2yz + y) * 2] + [(x^3z + x) * 8t] + [(x^3y + 1) * \frac{1}{t}]$$

Plug in x, y, and z in terms of t, and you have your answer in t.

Example. When you are asked to solve problems using these math tools, rarely you will be directly told to compute a certain derivative or integral. You have to use your own judgement about what needs to be done and translate that in math.

Suppose $y = f(x, w)$, while in turn $x = g(t, s)$ and $w = h(t, s)$. How does y change when t changes? When s changes? Let's derive the expressions to answer both of these questions.

$$\frac{\partial y}{\partial t} = \frac{\partial y}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial y}{\partial w}\frac{\partial w}{\partial t}$$

$$\frac{\partial y}{\partial s} = \frac{\partial y}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial y}{\partial w}\frac{\partial w}{\partial s}$$

You may recall from an example earlier where we introduced outputs $\boldsymbol{q}$ and inputs $\boldsymbol{z}$. You are given that $\boldsymbol{q} = f(\boldsymbol{z})$, which is a standard function denoting the level of production. Remember that both $\boldsymbol{q}$ and $\boldsymbol{z}$ are vectors that that there are two inputs $z_l and\ z_k$and one output q. Hint, it is very simply, do not worry about chain rule here.

As a practice, totally differentiate $\boldsymbol{q} = f(\boldsymbol{z})$ with respect to $z_l, z_k, and\ q$.

$$\frac{\partial q}{\partial q}dq = \frac{\partial f(\boldsymbol{z})}{\partial z_l}dz_l + \frac{\partial f(\boldsymbol{z})}{\partial z_k}dz_k$$
$$dq = \frac{\partial f(\boldsymbol{z})}{\partial z_l}dz_l + \frac{\partial f(\boldsymbol{z})}{\partial z_k}dz_k$$

3.6.3. Higher order derivatives

Think of one of the function's partial derivatives, for example, the partial with respect to x_1. We note that it is a function of n variables as follows:

$$\frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_1} = f_1(x)$$

If we were to calculate the n partial derivatives of $f_1(x)$, we get the *gradient* of second order derivatives, which is:

$$\nabla f(x) = \frac{\partial^2 f(x)}{\partial x_1 \partial x_1} + \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} + \dots + \frac{\partial^2 f(x)}{\partial x_n \partial x_1}$$

$$= f_{11}(x) + f_{12}(x) + \dots + f_{1n}(x)$$

A *Hessian Matrix* of the function $f(x)$ contains all the possible second-order partial derivatives of the original function.

$$D^2 f = Hessian\ (H) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f(x)}{\partial x_n \partial x_1} \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

Example. Let $y = e^{x_1} + 2x_2 + 4x_1 x_2^2$

$\mathcal{Q}$: What is the dimension of this Hessian for y?

$\mathcal{A}$: It is an $2\ x\ 2$ matrix

Let's find all that we need: $f_{11}, f_{12}, f_{21}, f_{22}$

$$f_1 = e^{x_1} + 4x_2^2; f_{11} = e^{x_1}; f_{12} = 8x_2$$
$$f_2 = 2 + 8x_1 x_2; f_{22} = 0; f_{21} = 8x_2$$

$$H = \begin{bmatrix} f_{11} & f_{21} \\ f_{12} & f_{22} \end{bmatrix}$$

$$= \begin{bmatrix} e^{x_1} & 8x_2 \\ 8x_2 & 0 \end{bmatrix}$$

Young's Theorem: Suppose that $f : \mathbb{R}^n \to \mathbb{R}$ is twice continuously differentiable in $\mathbb{R}^n$. Then, for all $x \in \mathbb{R}^n$ and for each pair of indices i, j:

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

Refer back to the example before, did we find that $f_{12} = f_{21}$?

$\mathcal{Q}$: Derive the Hessian for the function: $f(x, y) = (x^2 + y^2)e^{-y}$. Can you check the condition from Young's Theorem?

$\mathcal{A}$:

$f_x = 2xe^{-y}$	$f_{xx} = 2e^{-y}$
$f_y = e^{-y}(2y - x^2 - y^2)$	$f_{yy} = e^{-y}(2 - 4y + x^2 + y^2)$
$f_{xy} = -2xe^{-y}$	$f_{yx} = -2xe^{-y}$

As $f_{xy} = f_{yx} = -2xe^{-y}$, Young's Theorem holds.

The Hessian, $H = \begin{bmatrix} 2e^{-y} & -2xe^{-y} \\ -2xe^{-y} & e^{-y}(2 - 4y + x^2 + y^2) \end{bmatrix}$