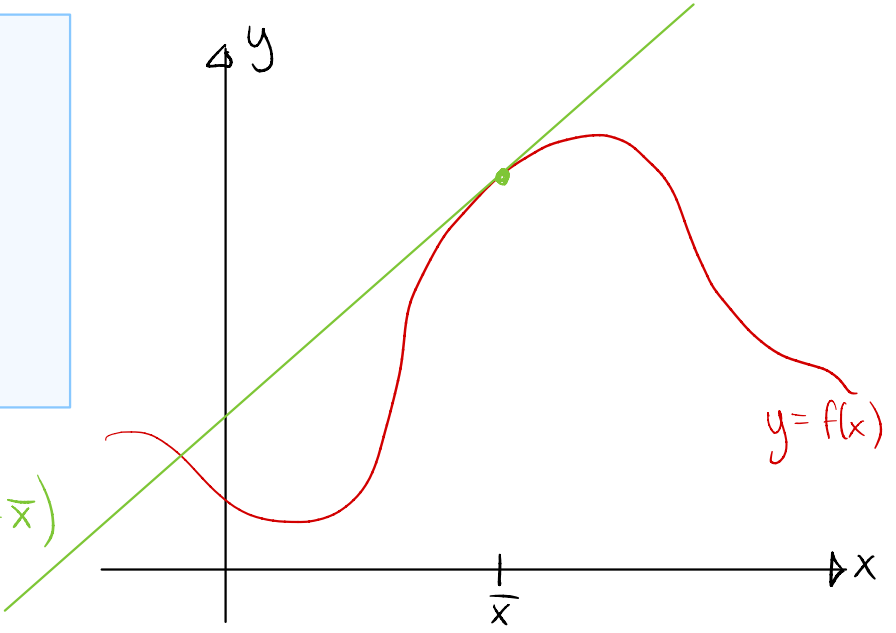
WHAT IS A DERIVATIVE?

Main questions in Chapter 6:

- What is the derivative of $F: U \to V$, when U, V are vector spaces?
- How do we compute it?
- What does it tell us about F?

The derivative of $f: \mathbb{R} \to \mathbb{R}$ at $\bar{x} \in \mathbb{R}$ is

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



$$y = f(\bar{x}) + f'(\bar{x})(x - \bar{x})$$

Geometrically, $f'(\bar{x})$ is the slope of the tangent of the graph of f at $x = \bar{x}$. But this is difficult/impossible to visualize in a general vector space!

For analysis purposes, the following is the most important property of f':

Taylor's theorem

Let $f: \mathbb{R} \to \mathbb{R}$ be a differentiable function.

Let $\bar{x} \in \mathbb{R}$. Then

$$f(\bar{x}+h) \approx f(\bar{x}) + f'(\bar{x})\,h \qquad \text{for small } h.$$

More precisely,

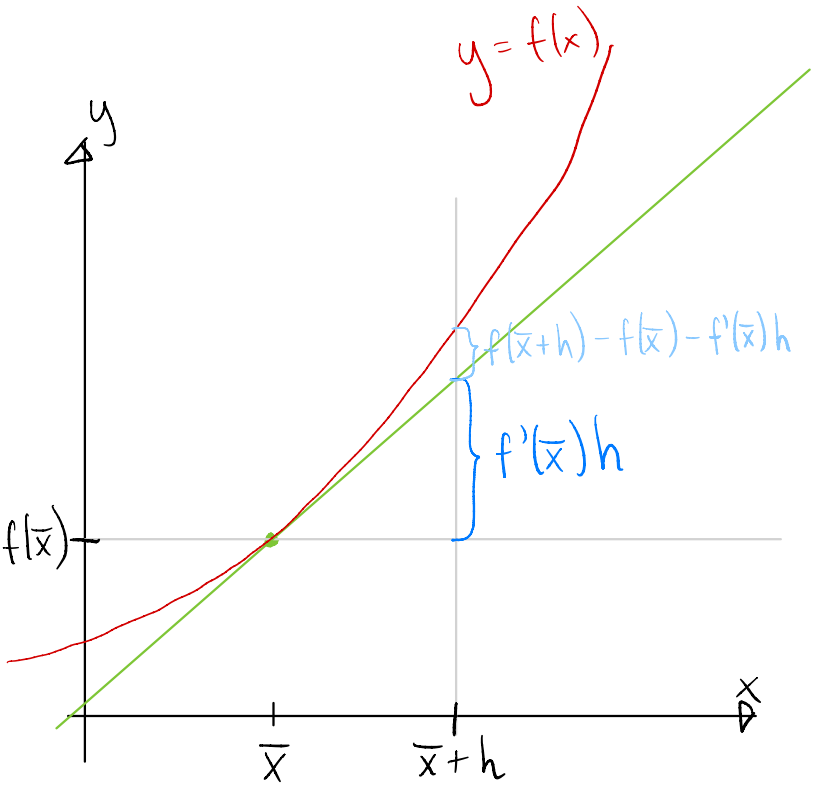
$$\frac{\left| f(\bar{x}+h) - \left(f(\bar{x}) + f'(\bar{x})\,h\right)\right|}{|h|} \xrightarrow[h \to 0]{} 0$$

In the expression

$$f(\bar{x}+h) \approx f(\bar{x}) + f'(\bar{x})\,h$$

$f'(\bar{x})$ does not appear on its own. It is <u>acting</u> on h.

$f'(\bar{x})\,h$ tells us (roughly) what f looks like at $\bar{x}+h$ (with larger accuracy for smaller h).



We interpret "$f(\bar{x}+h) \approx f(\bar{x}) + f'(\bar{x})h$ for small h"

as

$$\frac{|f(\bar{x}+h) - (f(\bar{x}) + f'(\bar{x})h)|}{|h|} \xrightarrow[h \to 0]{} 0.$$

This can be taken as a definition of $f'(\bar{x})$:

$$\frac{|f(\bar{x}+h) - (f(\bar{x}) + f'(\bar{x})h)|}{|h|} = \left| \frac{f(\bar{x}+h) - f(\bar{x})}{h} - f'(\bar{x}) \right| \xrightarrow[h \to 0]{!} 0$$

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

For a function $f: \mathbb{R}^m \to \mathbb{R}^n$ we identify the derivative with its <u>Jacobian</u> ∇f,

$$\nabla f(\bar{x}) = \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(\bar{x}) & \cdots & \frac{\partial f^1}{\partial x^m}(\bar{x}) \\ \vdots & & \\ \frac{\partial f^n}{\partial x^1}(\bar{x}) & \cdots & \frac{\partial f^n}{\partial x^m}(\bar{x}) \end{pmatrix} \in \mathbb{R}^{n \times m}$$

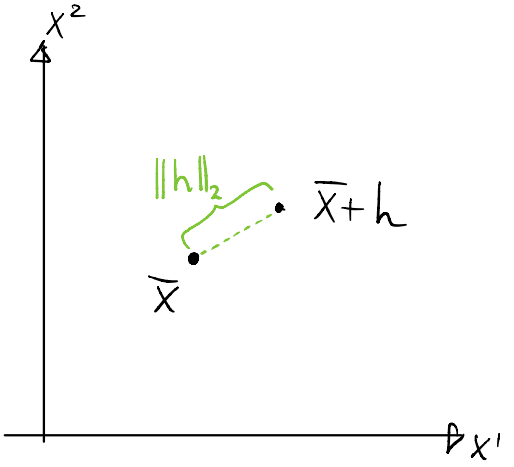
Then $f(\bar{x}+h) \approx f(\bar{x}) + \nabla f(\bar{x})\, h$ for h close to 0

More precisely, $$\frac{\| f(\bar{x}+h) - f(\bar{x}) - \nabla f(\bar{x})\, h \|_2}{\|h\|_2} \xrightarrow[h \to 0]{} 0$$

Note: In the expression

$$f(\bar{x}+h) \approx f(\bar{x}) + \nabla f(\bar{x})\, h$$

$\nabla f(\bar{x})$ does not appear on its own.
What is relevant is how it
acts on h.



We can think of $\nabla f(\bar{x})$
as a (linear) function of h.

We can now generalize to $F: X \to Y$ (where X, Y are normed vector spaces):

The derivative $F'(\bar{x})$ at $\bar{x} \in X$ should be some linear function from X to Y such that

$$F(\bar{x}+h) \approx F(\bar{x}) + F'(\bar{x})(h) \quad \text{for } h \text{ close to } 0.$$

QUESTIONS?

COMMENTS?

Next video: Define the Fréchet derivative.