Derivation of the parabola derivative

$$
\begin{aligned}
\lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} &= \lim_{h \to 0} \frac{x^2 + 2xh + h^2 - x^2}{h} && (4) \\
&= \lim_{h \to 0} \frac{2xh + h^2}{h} && (5) \\
&= \lim_{h \to 0} \frac{h(2x+h)}{h} && (6) \\
&= \lim_{h \to 0} 2x + h && (7) \\
&= 2x && (8)
\end{aligned}
$$

Derive on the board.

Derivate of a parabola:

$$
\begin{aligned}
\lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} &= \lim_{h \to 0} \frac{x^2 + 2xh + h^2 - x^2}{h} && (9) \\
&= \lim_{h \to 0} \frac{2xh + h^2}{h} && (10) \\
&= \lim_{h \to 0} \frac{h(2x+h)}{h} && (11) \\
&= \lim_{h \to 0} 2x + h && (12) \\
&= 2x && (13)
\end{aligned}
$$

The derivate of a polynomial

What is the derivative of the function $f(x) = x^n$?

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = nx^{n-1} \tag{14}$$

Derivate of a polynomial $f(x) = x^n$ [DFO20]:

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{(x+h)^n - x^n}{h} \tag{15}$$

$$= \lim_{h\to 0} \frac{\sum_{i=0}^{n} \binom{n}{i} x^{n-1} h^i - x^n}{h} \tag{16}$$

$$= \lim_{h\to 0} \frac{\sum_{i=1}^{n} \binom{n}{i} x^{n-1} h^i}{h} \tag{17}$$

$$= \lim_{h\to 0} \sum_{i=1}^{n} \binom{n}{i} x^{n-1} h^{i-1} \tag{18}$$

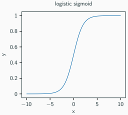
$$= \lim_{h\to 0} (\binom{n}{1} x^{n-1}) + \sum i = 2n \binom{n}{i} x^{n-i} h^{i-1} \tag{19}$$

$$= \frac{n!}{1!(n-1)!} x^{n-1} = nx^{n-1}. \tag{20}$$

The logistic sigmoid [GBC16]

The sigmoid function $\sigma(x)$ is a common activation function.

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{25}$$



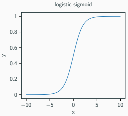
$$\begin{aligned}
\frac{\mathrm{d}\sigma(x)}{\mathrm{d}x} &= \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{1+e^{-x}} && (26)\\
&= \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{1+e^{-x}} \cdot 1 && (27)\\
&= \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{1+e^{-x}} \cdot \frac{e^x}{e^x} && (28)\\
&= \frac{\mathrm{d}}{\mathrm{d}x}\frac{e^x}{e^x+1} && (29)
\end{aligned}$$

$$g(x) = e^x, h(x) = e^x + 1 \tag{30}$$

The logistic sigmoid [GBC16]

The sigmoid function $\sigma(x)$ is a common activation function.

$$\sigma(x) = \frac{1}{1+e^{-x}} \tag{25}$$



$$\frac{\mathrm{d}\sigma(x)}{\mathrm{d}x} = \frac{e^x \cdot (e^x+1) - e^x \cdot e^x}{(e^x+1)^2} \tag{31}$$

$$= \frac{e^x \cdot e^x + e^x - e^x \cdot e^x}{(e^x+1)^2} \tag{32}$$

$$= \frac{e^x}{(e^x+1)^2} = \frac{e^x}{(e^x+1)}\frac{1}{(e^x+1)} \tag{33}$$

$$= \frac{e^x}{(e^x+1)}\left(\frac{1+e^x-e^x}{(e^x+1)}\right) \tag{34}$$

$$= \frac{e^x}{(e^x+1)}\left(\frac{1+e^x}{(e^x+1)} - \frac{e^x}{(e^x+1)}\right) \tag{35}$$

$$= \frac{e^x}{(e^x+1)}\left(1 - \frac{e^x}{(e^x+1)}\right) \tag{36}$$

$$= \sigma(x)(1-\sigma(x)) \tag{37}$$

The gradient

The gradient lists partial derivatives with respect to all inputs in a vector. For a function $f : \mathbb{R}^n \to \mathbb{R}$ of n variables the gradient $\nabla f : \mathbb{R}^n \to \mathbb{R}^n$ is defined as

$$\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}. \tag{45}$$

- Gradients point in the steepest ascent direction.
- To find the gradient, we must compute the partial derivate with respect to every input.
- A vector collects all derivates.

The gradient of the Rosenbrock function

Recall the Rosenbrock function:

$$f(x, y) = (a - x)^2 + b(y - x^2)^2 \tag{51}$$

$$\nabla f(x, y) = \begin{pmatrix} -2a + 2x - 4byx + 4bx^3 \\ 2by - 2bx^2 \end{pmatrix} \tag{52}$$

On the board, derive:

$$\begin{aligned}
f(x, y) &= (a - x)^2 + b(y - x^2)^2 && (53)\\
&= a^2 - 2ax + x^2 + b(y^2 - 2yx^2 + x^4) && (54)\\
&= a^2 - 2ax + x^2 + by^2 - 2byx^2 + bx^4 && (55)\\
\Rightarrow \frac{\partial f(x, y)}{\partial x} &= -2a + 2x - 4byx + 4bx^3 && (56)\\
\Rightarrow \frac{\partial f(x, y)}{\partial y} &= 2by - 2bx^2 && (57)
\end{aligned}$$

Temporary page!

LaTeX was unable to guess the total number of pages correctly. As there was some unprocessed data that should have been added to the final page this extra page has been added to receive it.

If you rerun the document (without altering it) this surplus page will go away, because LaTeX now knows how many pages to expect for this document.