Gradient and backpropagation

Gradient

Derivation function approximation

The definition of derivation is

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

With Δx is Infinitely small value. If we substitute it as a "small value", it would not cause a big offset. We get the following equation

$$f(x + \Delta x) \approx \frac{f(x + \Delta x) - f(x)}{\Delta x} \approx f'(x)$$

$$so, f'(x)\Delta x + f(x) \approx f(x + \Delta x)$$

e.g. If we want the get the approximation of e^x when x is close to 0, use the above equation we get

$$e^{x+\Delta x} \approx e^x + e^x \Delta x$$

Then we replace x with 0, $\Delta x\ with\ x$, we get $e^x \approx 1 + x$ when x is close to 0.

Gradient

Lagrange multiplier (Lagrangian function).

https://www.zhihu.com/question/38586401

https://en.wikipedia.org/wiki/Lagrange_multiplier

The method can be summarized as follows: in order to find the maximum or minimum of a function $f(x)$ subject to the equality constraint $g(x) = 0$, form the Lagrangian function

$$L(x, \lambda) = f(x) - \lambda g(x)$$

And find the stationary points of L considered as a function of x and the Lagrange multiplier λ