# Gradient and backpropagation

## Gradient

### Derivation function approximation

The definition of derivation is

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

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

Then we replace with 0, , we get when x is close to 0.

### Gradient

The value of the gradient at a point is a **tangent vector**. The gradient vector can be interpreted as the **“direction and rate of fastest increase”.** if the gradient of is defined as follows:

### Lagrange multiplier (Lagrangian function).

The method can be summarized as follows: in order to find the maximum or minimum of a function subject to the equality constraint , form the Lagrangian function

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

How to understand it?

Imagine we want to calculate the closest distance from to the origin. Here is one solution: We draw a circle with function , when the circle first time tangent with our blue curve, the radius of the circle is the closest distance.

Diagram, schematic, radar chart

Description automatically generated

When the circle first time tangent with the blue curve, their normal vector (gradient) should be parallel with each other, which written as

Chart

Description automatically generated

So, for now, we get two constraints,

Calculate the gradient and put them into this formula, we get

Then we get