Use of Derivatives

MATHEMATIC

❖ Differential equation

Coefficients and degrees

- The coefficient of a differential equation is the largest number of differentials
- For example, the next red determines the coefficient of the differential equation

$$x^3 \frac{dy}{dx} + 2y = 0, \quad \frac{d^2 y}{dx^2} + 2x^2 \frac{dy}{dx} = 0$$

- The red term in the equation on the left is the same as $\frac{dy^1}{dx^1}$, so the coefficient is 1
- First-order differential equation
- The formula on the right is called a quadratic differential equation because the largest number of differentials is 2

Differentiation

❖ Differential equation

- The degree is the number of powers of the highest coefficient term of the differential equation
- For example, the next red term determines the order of the differential equation

$$x^3\left(\frac{dy}{dx}\right)^1 + 2y = 0, \quad \left(\frac{d^2y}{dx^2}\right)^2 + 2x^2\frac{dy}{dx} = 0$$

- The red term in the equation on the left is 1, so it is called a first-order differential equation
- The formula on the right is called a second-order quadratic differential equation because the number of powers in the highest coefficient term is 2

Differentiation

❖ **Differential equation**

- If the coefficients in each term depend only on the independent variable x, it is called ***linear***
- If any of the coefficients has a term that depends on the dependent variable y, it is called ***non-linear***

$$x\frac{dy}{dx}$$

선형 예시

$$y\frac{dy}{dx}$$

비선형 예시

❖ Differential equation

Ordinary differential equation and partial differential equation

- Equations containing unknown derivatives are called differential equations
- There are two differential equations: ordinary differential equation (ODE) and a partial differential equation (PDE)
- The commonly used differentiation method is typically to obtain $\frac{dy}{dx}$
- In other words, if **only the derivative differentiated by one independent variable** is included, it is a ODE
- If derivatives that are differentiated by **more than one independent variable** is included,

  it is PDE
- Independent variable:
  - Equation of y = f(x) is established when the value of the function y also changes as the value of the x variable changes
  - At this time, x is called the independent variable, and y is called the dependent variable

Differentiation

❖ Differential equation

- All equations covered so far are ODE such as

$$y = x^2 + 1$$

$$y = 2x^2 + 4x + 3$$

MATHEMATIC

❖ Differential equation

- PDE is when you have two variables, x and y, and you see one of them as a constant
- f the variable is x and y, there are two cases: $\frac{dy}{dx}, \frac{dx}{dy}$
- That is, there are two ways to differentiate x (differentiate y for x) by looking at y as a constant, and to differentiate y (differentiate x for y) by looking at x as a constant

MATHEMATIC

❖ Differential equation

- PDE symbols for function z = f(x, y) are

$$x\text{에 대한 편미분: } f_x(x,y) = f_x = \frac{\partial f}{\partial x} = \frac{\partial}{\partial x} f(x,y) = \frac{\partial z}{\partial x}$$

$$y\text{에 대한 편미분: } f_y(x,y) = f_y = \frac{\partial f}{\partial y} = \frac{\partial}{\partial y} f(x,y) = \frac{\partial z}{\partial y}$$

Differentiation

❖ Differential equation

- The method of calculating PDE is simple as follows
  - When finding f_x(x, y), see y as a constant and differentiate f(x, y) with respect to x
  - When f_y(x, y), see x as a constant and differentiate f(x, y) with respect to y

❖ Differential equation

- For example, let’s partial differentiate the function z = $2x^2 + 3x^2y - 5xy + 3x$
  - (1) y as a constant and differentiate for x

$$\frac{\partial z}{\partial x} = 4x\,dx + 6xy\,dx - 5y\,dx + 3\,dx$$

$$= (4x + 6xy - 5y + 3)dx$$

$$\therefore \frac{\partial z}{\partial x} = 4x + 6xy - 5y + 3$$

MATHEMATIC

❖ Differential equation

- (2) x as a constant and differentiate for y

$$\frac{\partial z}{\partial y} = 3x^2\,dy - 5x\,dy$$

$$= (3x^2 - 5x)dy$$

$$\therefore \frac{\partial z}{\partial y} = 3x^2 - 5x$$

❖ Differential equation

연습 문제

$f(x, y) = 3x^2 + 2y^2 + xy$일 때 $f_x(2, 3)$과 $f_y(2, 3)$의 값을 구하세요.

❖ Differential equation

연습 문제

$f(x, y) = 3x^2 + 2y^2 + xy$일 때 $f_x(2, 3)$과 $f_y(2, 3)$의 값을 구하세요.

문제 풀이

(1) $f_x(x, y)$에서 y를 상수로 보고 x에 대해 미분합니다.

$f_x(x, y) = 6x + y$이므로 $f_x(2, 3) = 6 \times 2 + 3 = 15$입니다.

(2) $f_y(x, y)$에서 x를 상수로 보고 y에 대해 미분합니다.

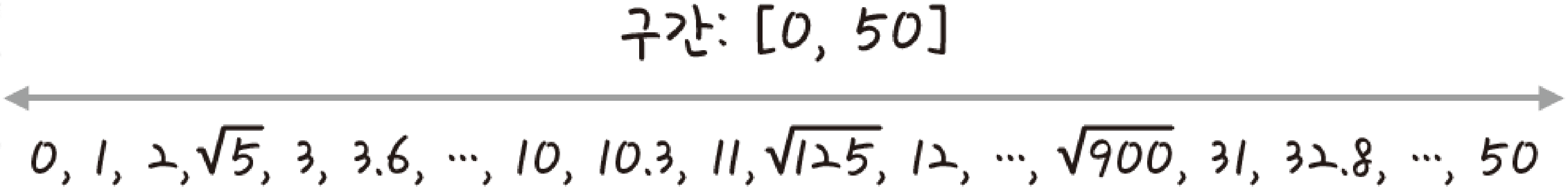
$f_y(x, y) = 4y + x$이므로 $f_y(2, 3) = 4 \times 3 + 2 = 14$입니다.

MATHEMATIC

❖ Mean value theorem and Rolle's theorem

Interval

- In general, an interval means between a point and another point
- For example, if a section of the road is controlled, it means that the section is controlled from road A to road B
- In the mathematical, an interval is the set of all real numbers that lie between two real numbers on the vertical line
- For example, if there are 0 to 50 real numbers on the vertical line, the interval between all the real numbers



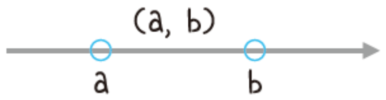
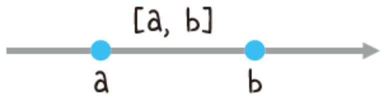
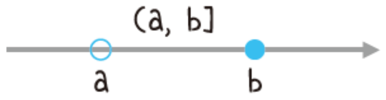
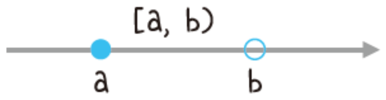
MATHEMATIC

❖ Mean value theorem and Rolle’s theorem

- The symbol for the interval is as follows

$$(a,\ b),\ [a,\ b],\ (a,\ b],\ [a,\ b)$$

- A closed interval contains both ends in the set of real numbers, and an open interval does not contain both ends in the set of real numbers

기호	의미		용어
(a, b)		$a < x < b$	열린 구간(open interval), 개구간
[a, b]		$a \leq x \leq b$	닫힌 구간(closed interval), 폐구간
(a, b]		$a < x \leq b$	반열린 구간(half–open interval)
[a, b)		$a \leq x < b$	

❖ Mean value theorem and Rolle's theorem

Mean value theorem

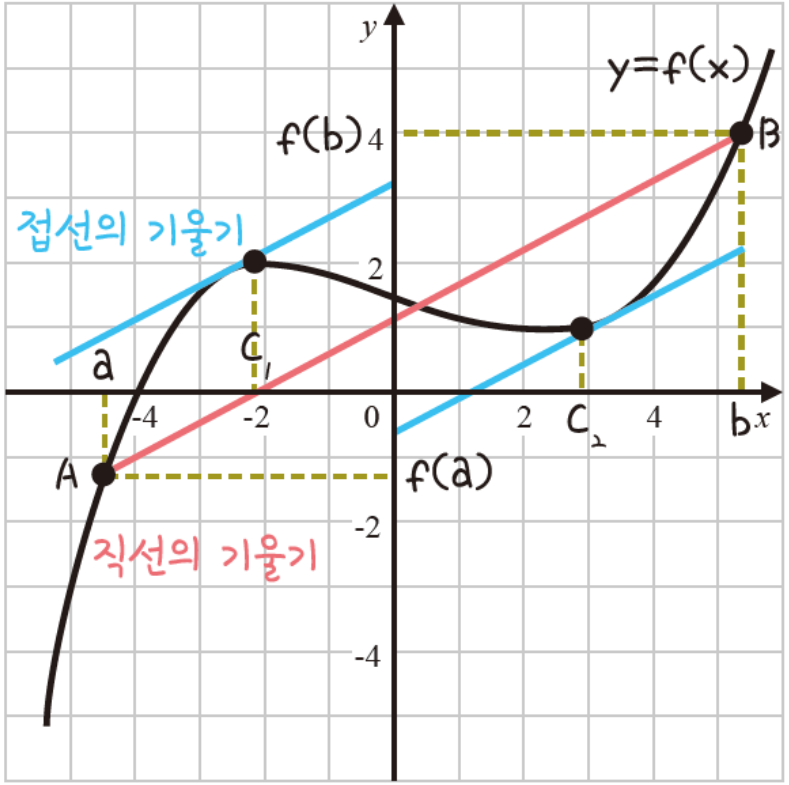
- Mean value theorem can be defined as follows

함수 $f(x)$가 닫힌 구간 $[a, b]$에서 연속이고 열린 구간 (a, b)에서 미분 가능하면
$\dfrac{f(b)-f(a)}{b-a}=f'(c)$인 c가 a와 b 사이에 적어도 하나 존재합니다.

- $\frac{f(b)-f(a)}{b-a}$ is the average rate of change when function f(x) is given in the interval [a, b]
- f'(c) is the derivative at x = c in the presence of f(x)
- The average rate of change is the slope of straight line passing through any two points, and the derivative is the slope of the tangent line

❖ Mean value theorem and Rolle's theorem

- $\frac{f(b)-f(a)}{b-a} = f'(c)$ means that there is a slope of the tangent lines (c1 and c2) such as the slope of the straight line passing through the two vertices A and B



❖ Mean value theorem and Rolle's theorem

- In other words, the mean value theorem means that among the tangents of the curve y = f(x), there is at least one contact (c1 and c2) parallel to the straight lines A and B between a and b
- Sometimes the mean value theorem can be used to prove it simply
- In other words, it is used to simply prove calculus that is difficult to graph

❖ Mean value theorem and Rolle’s theorem

연습 문제

함수 $f(x)=x^2$에서 구간 $[0,\ 5]$에서 평균값 정리를 만족하는 c 상수 값을 구하세요.

MATHEMATIC

❖ Mean value theorem and Rolle's theorem

연습 문제

함수 $f(x)=x^2$에서 구간 $[0,\ 5]$에서 평균값 정리를 만족하는 c 상수 값을 구하세요.

문제 풀이

함수 $f(x)$는 닫힌 구간 $[0,\ 5]$에서 연속이고 열린 구간 $(0,\ 5)$에서 미분 가능합니다.

평균값 정리에 따라 $\dfrac{f(5)-f(0)}{5-0}=f'(c)$를 만족하는 c가 구간 $(0,\ 5)$에서 적어도 하나가 존재해야 합니다. 따라서 $f'(c)=\dfrac{f(5)-f(0)}{5-0}=\dfrac{25-0}{5-0}=5$이고 $f'(x)=2x$이므로, $f'(c)=2c=5$가 되어 $c=2.5$입니다.

❖ Mean value theorem and Rolle's theorem

Rolle's theorem

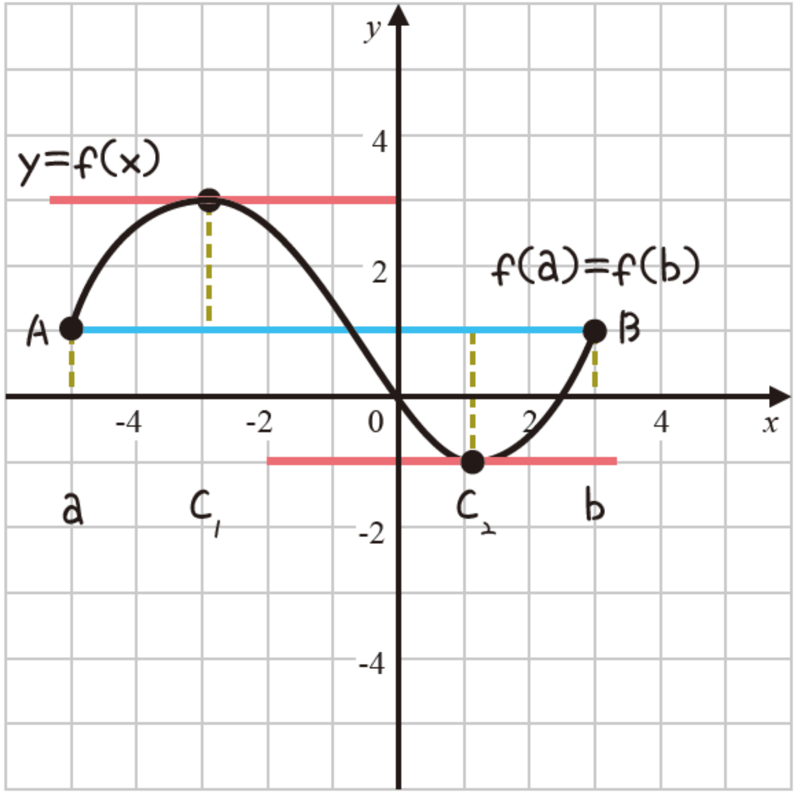
- Rolle's theorem is as follows

함수 $f(x)$가 닫힌 구간 $[a,\ b]$에서 연속이고 열린 구간 $(a,\ b)$에서 미분 가능할 때, $f(a)=f(b)$이면 $f'(c)=0\ (a<c<b)$인 c가 열린 구간 $(a,\ b)$에 적어도 하나 존재합니다.

MATHEMATIC

❖ Mean value theorem and Rolle's theorem

- Given a function y = f(x), f(a) = f(b) is a curve as shown below (but continuous on closed sections [a, b], differentiable on open sections (a, b))



MATHEMATIC

❖ Mean value theorem and Rolle's theorem

- Rolle's theorem is a special case of the mean value theorem, and you can think of the average rate of change as zero

연습 문제

$f(x) = 2x^3 - 6x^2$ $[0, 5]$ 함수에서 롤의 정리를 만족하는 c 값을 구하세요.

❖ Mean value theorem and Rolle's theorem

- Rolle's theorem is a special case of the mean value theorem, and you can think of the average rate of change as zero

연습 문제

$f(x) = 2x^3 - 6x^2$ $[0,\ 5]$ 함수에서 롤의 정리를 만족하는 c 값을 구하세요.

문제 풀이

함수 $f(x)$는 닫힌 구간 $[0,\ 5]$에서 연속이고 열린 구간 $(0,\ 5)$에서 미분 가능합니다. $f(0) = f(5) = 0$이므로 $f'(c) = 0$을 만족하는 c가 $0 < c < 5$에서 적어도 하나 존재합니다.

MATHEMATIC

❖ Mean value theorem and Rolle's theorem

$$f(x) = 2x^3 - 6x^2$$

$$f'(x) = (2x^3 - 6x^2)'$$

$$f'(c) = (2c^3 - 6c^2)'$$

$$= 6c^2 - 12c$$

$$= 6c(c-2)$$

$$\therefore\ c = 0,\ 2$$

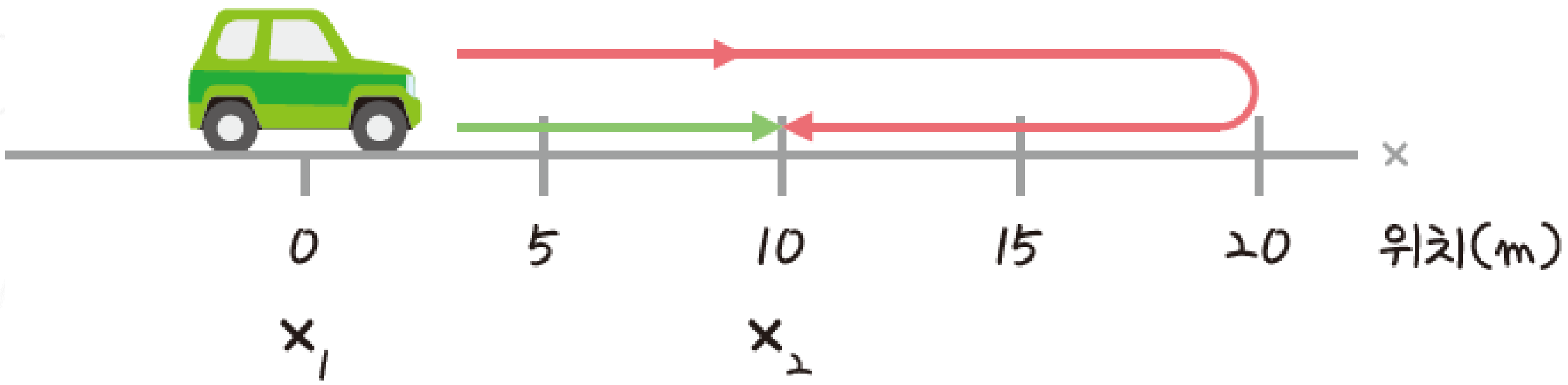
즉, $0 < c < 5$이므로 $c = 2$입니다.

MATHEMATIC

❖ Velocity and acceleration

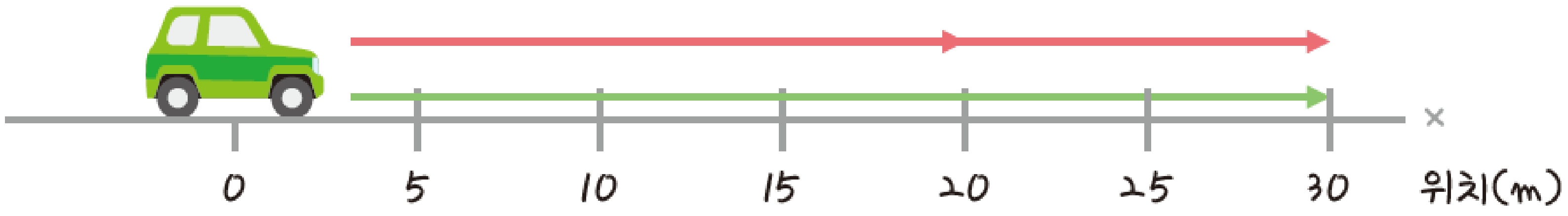
Distance and displacement

- The distance is the distance an object has traveled
- If the car moved from point A to point B, 100 meters away, the distance by the car is 100 meters
- Distance is an undirected, size-only physical quantity, also known as a scalar quantity
- After moving up to 20m, it's back 10m

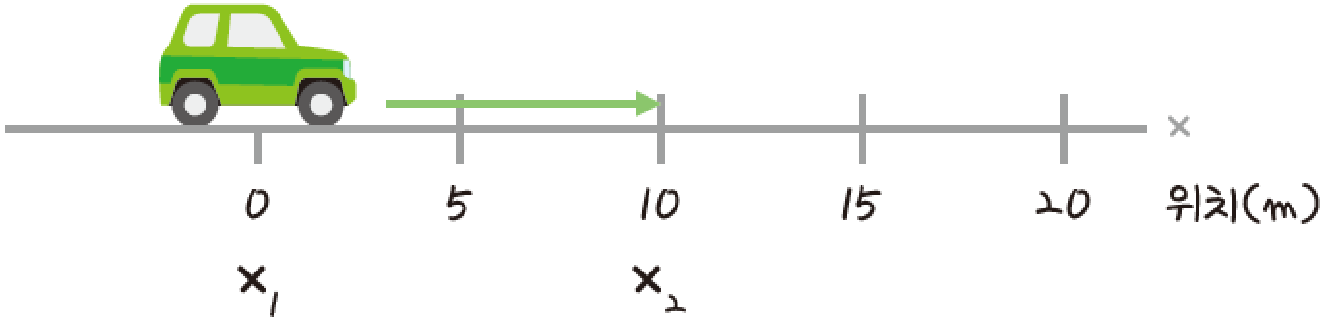


❖ Velocity and acceleration

- The distance is said to have no direction, so the travel distance is a total of 30m as follows



- Displacement is not only in size but also in direction because it represents how far away the point of arrival is from the starting point, i.e. a change in position
- These physical quantities are called vector quantities, and the displacement of the car is 10 m

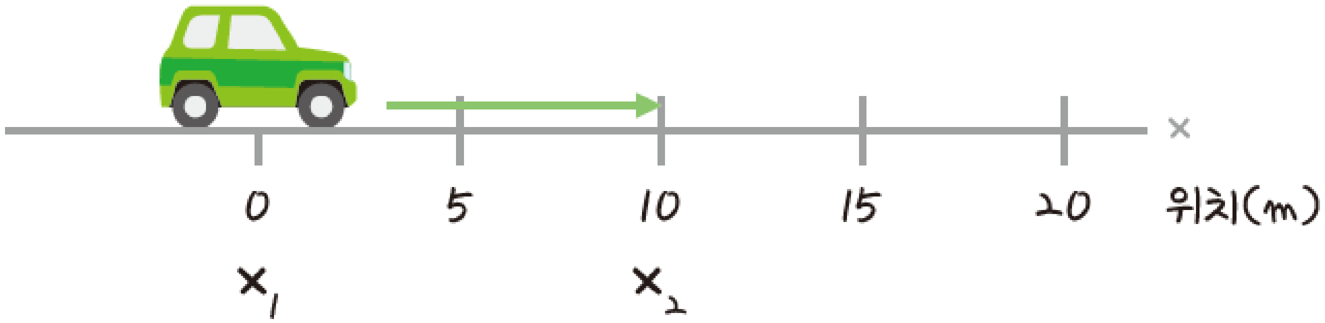


❖ Velocity and acceleration

- The displacement is expressed in a formula

$$\Delta x = x_2 - x_1$$

- Because the displacement is defined as Δx = x2 - x1, the car displacement in previous Figure is +10m - 0m = 10m (displacement is also said to have direction, so +, - representation is also required)



❖ Velocity and acceleration

Speed and velocity

- Learning distance and displacement from differential is because distance and displacement are necessary to calculate speed and velocity
- Let's look at the formula for calculating speed and velocity

$$\text{속력} = \frac{\text{이동 거리}}{\text{시간}}$$

$$\text{속도} = \frac{\text{변위}}{\text{시간}}$$

Differentiation

❖ Velocity and acceleration

- In other words, speed does not take into account the direction of movement, but velocity does take into account the direction of movement
- Assuming that it takes 1 second for the car to go 5m, the speed and velocity are calculated as follows

$$속력 = \frac{30\text{m}}{6초} = 5\text{m/s}$$

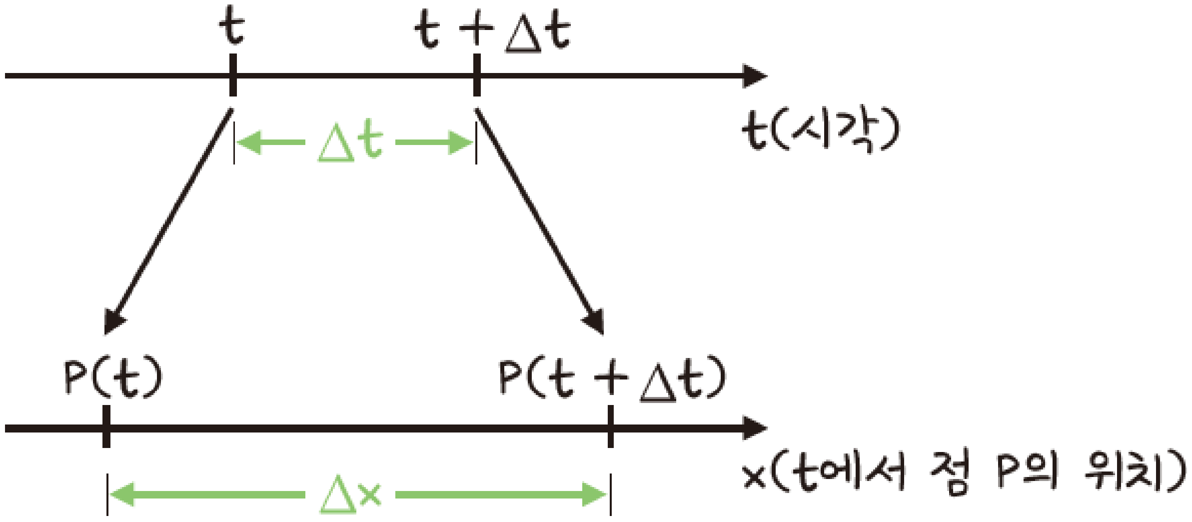
$$속도 = \frac{10\text{m}}{6초} = 약\ 1.7\text{m/s}$$

Differentiation

❖ Velocity and acceleration

Velocity and acceleration

- Average rate of change is the rate of change between coordinate points A and B
- The instantaneous rate of change is the rate of change at a specific point in one moment
- **In differential, velocity means instantaneous rate of change, not average rate of change**
- In other words, when a position is expressed as a function of time, differentiating the position over time yields velocity



MATHEMATIC

❖ Velocity and acceleration

- In other words, if the position of the point P at time t is expressed in coordinate x when P moves over the vertical line, x can be expressed as a function x = f(t) for t
- At this time, when the time changes from t to t + Δt, the average velocity is equal to the average rate of change of the function f(t), so it can be expressed by the following formula

$$\frac{\Delta x}{\Delta t} = \frac{f(t+\Delta t) - f(t)}{\Delta t}$$

❖ Velocity and acceleration

- Similarly, at the time t of function f(t), the instantaneous rate of change is as follows

$$\lim_{\Delta t \to 0} \frac{\Delta x}{\Delta t} = \frac{dx}{dt}$$

- The instantaneous rate of change of the function x = f(t) is called the instantaneous velocity of point P, and the symbol is expressed as v
- This means that the instantaneous velocity v is

$$v = \frac{dx}{dt} = \lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}$$

❖ Velocity and acceleration

- Given velocity v at time t for a point P moving over a vertical line, the derivative $\frac{dv}{dt}$ of velocity v represents the instantaneous rate of change of velocity at time t for a point P
- In this way, the change in velocity over time t, i.e., the instantaneous rate of change in velocity v, is called acceleration and the symbol a is used

$$a = \lim_{\Delta t \to 0} \frac{\Delta v}{\Delta t} = \frac{dv}{dt}$$

❖ Velocity and acceleration

연습 문제

원점과 수직선 위 임의의 점 P의 시간 t에서 위치를 $x = 2t^3 - 6t$라고 할 때 다음 질문에 답하세요.

(1) $t = 3$에서 점 P의 속도와 가속도를 구하세요.

(2) 점 P의 진행 방향이 바뀌는 시각을 구하세요.

❖ Velocity and acceleration

문제 풀이

속도와 가속도 정의에 따라 $x=2t^3-6t$는 $v=6t^2-6$, $a=12t$가 됩니다.

(1) $t=3$일 때

$v=6\times 3^2-6=6\times 9-6=48$

$a=12\times 3=36$

따라서 속도는 48, 가속도는 36입니다.

(2) 점 P의 진행 방향이 바뀌는 시각에서 $v=0$이므로

$6t^2-6=0$에서 $6t^2=6$이 되어 $t=1$입니다($\because\ t>0$).

따라서 점 P의 진행 방향이 바뀌는 시각은 $t=1$입니다.

Differentiation and backpropagation

❖ Differential rules

- Differential rules for more than one function with the same variable include ***sum(difference) rule***, ***product rule***, ***quotient rule***, and ***chain rule***

❖ Sum(difference) rule

- If both f(x) and g(x) are differentiable, the sum (difference) derivative of the functions is equal to the sum (difference) of each derivative

$$\frac{d}{dx}[f(x) \pm g(x)] = \frac{d}{dx} f(x) \pm \frac{d}{dx} g(x) = f'(x) \pm g'(x)$$

MATHEMATIC

❖ Differential rules

연습 문제

함수 $y = x^2 - 5x$를 합 법칙을 적용하여 미분하세요.

❖ Differential rules

연습 문제

함수 $y = x^2 - 5x$를 합 법칙을 적용하여 미분하세요.

문제 풀이

$$f(x) = x^2 - 5x$$

$$\begin{aligned} f'(x) &= (x^2)' + (-5x)' \\ &= (2x) + (-5) \\ &= 2x - 5 \end{aligned}$$

Differentiation and backpropagation

❖ Differential rules

Product rule

- If both f(x) and g(x) can be differentiated, it is the same as adding 'the product of the derivative of f(x) and g(x)' and 'the product of the function of f(x) and g(x)'

$$\frac{d}{dx}[f(x)g(x)] = f(x)g'(x) + f'(x)g(x)$$

Differentiation and backpropagation

❖ Differential rules

연습 문제

함수 $f(x) = (3x^2 - 1)(2x^2 + 4x + 2)$에 곱 법칙을 적용하여 미분하세요.

Differentiation and backpropagation

MATHEMATIC

❖ Differential rules

연습 문제

함수 $f(x) = (3x^2 - 1)(2x^2 + 4x + 2)$에 곱 법칙을 적용하여 미분하세요.

문제 풀이

$$
\begin{aligned}
f'(x) &= (3x^2 - 1)(2x^2 + 4x + 2)' + (3x^2 - 1)'(2x^2 + 4x + 2) \\
&= (3x^2 - 1)(4x + 4) + 6x(2x^2 + 4x + 2) \\
&= (12x^3 + 12x^2 - 4x - 4) + (12x^3 + 24x^2 + 12x) \\
&= 24x^3 + 36x^2 + 8x - 4
\end{aligned}
$$

Differentiation and backpropagation

❖ Differential rules

Quotient rule

- If both f(x) and g(x) can be differentiated, the derivative of the numerator is multiplied by the denominator, and the numerator is subtracted by the derivative of the denominator divided by the square of the denominator

$$\frac{d}{dx}[\frac{f(x)}{g(x)}] = \frac{g(x)[f(x)'] - f(x)[g(x)']}{[g(x)]^2} = \frac{f'g - g'f}{g^2}$$

Differentiation and backpropagation

❖ Differential rules

연습 문제

함수 $f(x) = \dfrac{2x^2 + 3x}{x^2 + 1}$ 의 도함수 $f'(x)$를 구하세요.

Differentiation and backpropagation

❖ Differential rules

연습 문제

함수 $f(x) = \dfrac{2x^2+3x}{x^2+1}$ 의 도함수 $f'(x)$를 구하세요.

문제 풀이

$$\begin{aligned} f'(x) &= \frac{(x^2+1)(2x^2+3x)' - (2x^2+3x)(x^2+1)'}{(x^2+1)^2} \\ &= \frac{(x^2+1)(4x+3) - (2x^2+3x)2x}{(x^2+1)^2} \\ &= \frac{(4x^3+3x^2+4x+3) - (4x^3+6x^2)}{(x^2+1)^2} \\ &= \frac{-3x^2+4x+3}{(x^2+1)^2} \end{aligned}$$

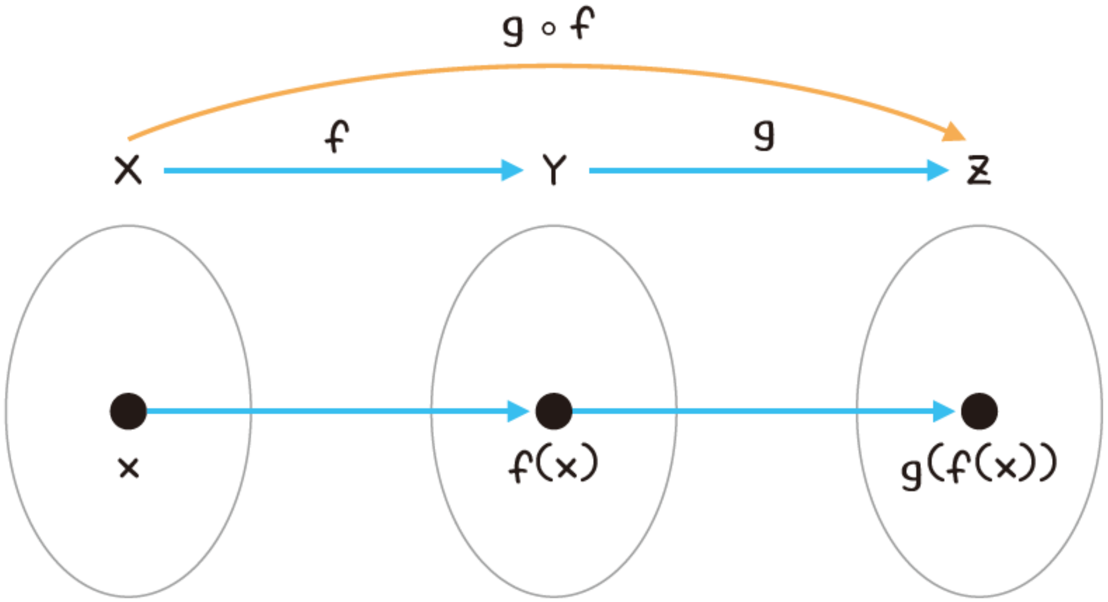
Differentiation and backpropagation

MATHEMATIC

❖ Differential rules

Chain rule

- The Chain rule is a differential method of a composite function that synthesizes two functions
- Given two functions f: X → Y, g: Y → Z, we can create a function from X to Z by corresponding f(x) to any element x in the set X and corresponding f(x) to g(f(x))



Differentiation and backpropagation

❖ Differential rules

Chain rule

- (1) the chain rule of a single variable function (a function with one variable)
  - If the single variable composite function y = f (g(x)) is decomposed by y = f (u), u = g (x), the following formula holds

$$y = f(u),\ u = g(x)$$

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

$$\Leftrightarrow \frac{d}{dx} f(u) = \frac{d}{du} f(u) \cdot \frac{d}{dx} g(x)$$

$$\Leftrightarrow \frac{d}{dx} f(g(x)) = f'(u) \cdot g'(x)$$

$$\Leftrightarrow [f(g(x))]' = f'(g(x)) \cdot g'(x)$$

Differentiation and backpropagation

MATHEMATIC

❖ Differential rules

- In differentiating dy and dx in each term, the term du is called a chain rule because one side is connected like a denominator and the other is a numerator chain
- Let's look at the chain rule of a function of single variable as an example
- For example, differentiate the function $y = (3x^2 + 2)^2$
- If $3x^2 + 2$ is a, then $a = 3x^2 + 2$ is a composite function of the two functions,
  $y = a^2$, $a = 3x^2 + 2$
- If the chain rule is applied at this time, it is as follows

$$\begin{aligned}\frac{dy}{dx} &= 2a^{2-1} \times ((3 \times 2)x^{2-1} + 2^0) \\ &= 2(3x^2 + 2) \times (6x + 0) \quad \longleftarrow \quad a\text{에 } 3x^2 + 2 \text{ 대입} \\ &= (6x^2 + 4) \times 6x \\ &= 36x^3 + 24x\end{aligned}$$

Differentiation and backpropagation

❖ Differential rules

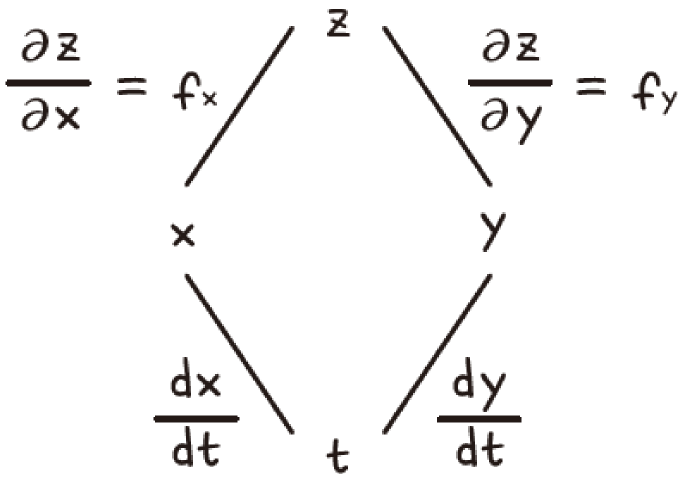
- (2) Chain rule of multivariate functions (functions with more than one variable)
  - The chain law of a single variable function can also be applied to a multivariate function
  - For example, if x = g(t), y = h(t), and f(x, y), and f(x, y), g(t), and h(t) are differentiable for a two-variable function z = f(x, y), the following formula holds

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

Differentiation and backpropagation

❖ Differential rules

- The more multivariate, the more difficult the chain rule is, and it is easier to understand it in order (1) to (3) by referring to the following



(1) z를 x에 대해 편미분한 후 x를 t에 대해 미분

(2) z를 y에 대해 편미분한 후 y를 t에 대해 미분

(3) z를 t에 대해 미분 = (1)과 (2)의 합

Differentiation and backpropagation

❖ Differential rules

연습 문제

함수 $y = \dfrac{1}{(2x^2+3x-1)^3}$ 에 대해 미분하세요.

Differentiation and backpropagation

MATHEMATIC

❖ **Differential rules**

연습 문제

함수 $y = \dfrac{1}{(2x^2+3x-1)^3}$ 에 대해 미분하세요.

문제 풀이

$y = \dfrac{1}{(2x^2+3x-1)^3}$ 에서 $2x^2+3x-1$을 z라고 한다면 $z=2x^2+3x-1$, $y=z^{-3}$처럼 z와 y 값을 얻을 수 있습니다.

이때 연쇄 법칙을 적용하면, $\dfrac{dy}{dx} = \dfrac{dz}{dx} \cdot \dfrac{dy}{dz}$ 이므로 미분 결과는 다음과 같습니다.

$$((2 \times 2)x^{2-1}+3)(-3z^{-3-1})$$

$$= -3(4x+3)(2x^2+3x-1)^{-4}$$

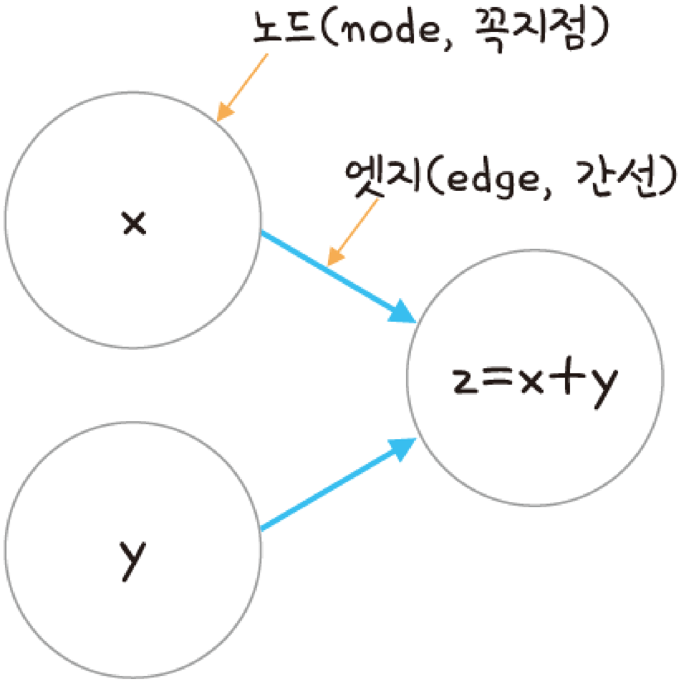
$$= \frac{-3(4x+3)}{(2x^2+3x-1)^4}$$

Differentiation and backpropagation

❖ Backpropagation

Computational graph

- A computational graph is a graph of the calculation process and is expressed in nodes and edges
- The node defines the operation, and the edge indicates the direction in which the data flows



Differentiation and backpropagation

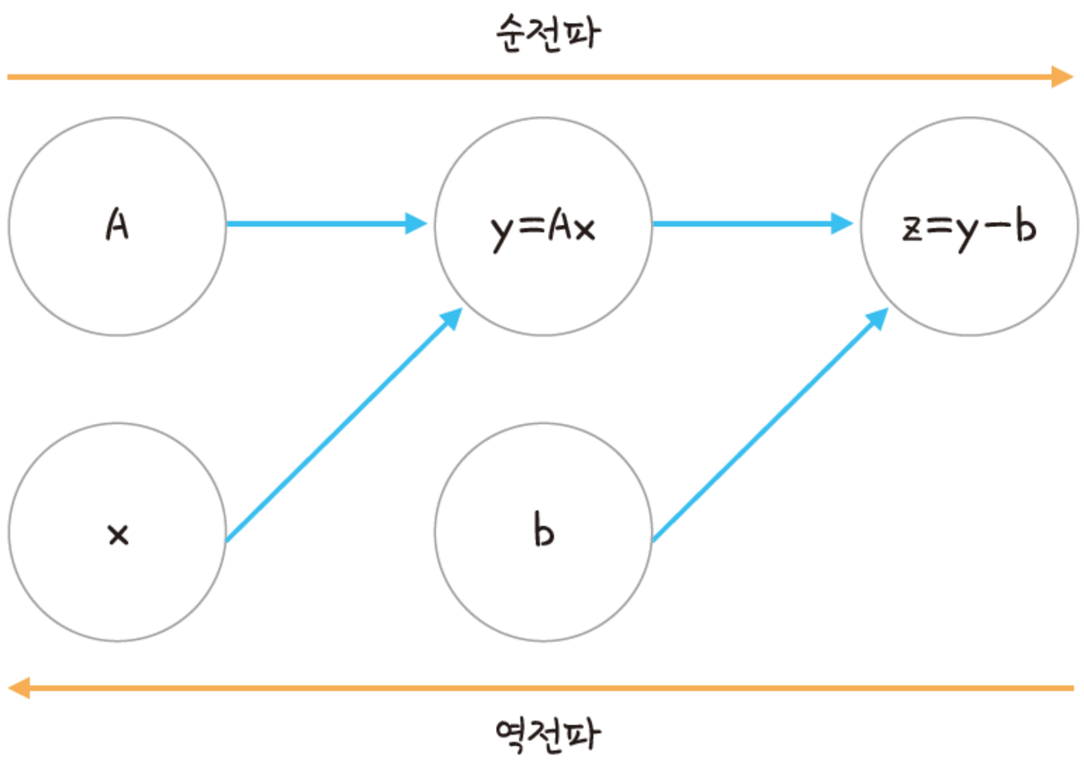
❖ Backpropagation

- The use of computational graphs in the computational process is due to the advantages
  - Local computation can simplify the problem by focusing on the computation of each node
    - Local computation: means calculating only within the scope of a calculation graph that is directly related to you
  - Can efficiently calculate differentials with backpropagation

Differentiation and backpropagation

❖ Backpropagation

- There are two methods for solving computational graphs: forward propagation and backward propagation
- If the calculation graph progresses from left to right, it is called a net wave
- On the other hand, if you go from right to left, it's called backpropagation



Differentiation and backpropagation

❖ **Backpropagation**

- Backpropagation corrects the **weight and bias of node values involved in this error by obtaining an error between the calculation result and the correct answer**
- The backpropagation is repeatedly corrected **in the process of decreasing the error**
- As the process repeated, the accuracy increases, but there is a disadvantage that it takes a long time
- Fewer times reduce accuracy, but reduce time
- This number of cycles is called an **epoch**
- **Increasing the epoch, updating (learning) weights and biases to reduce errors**

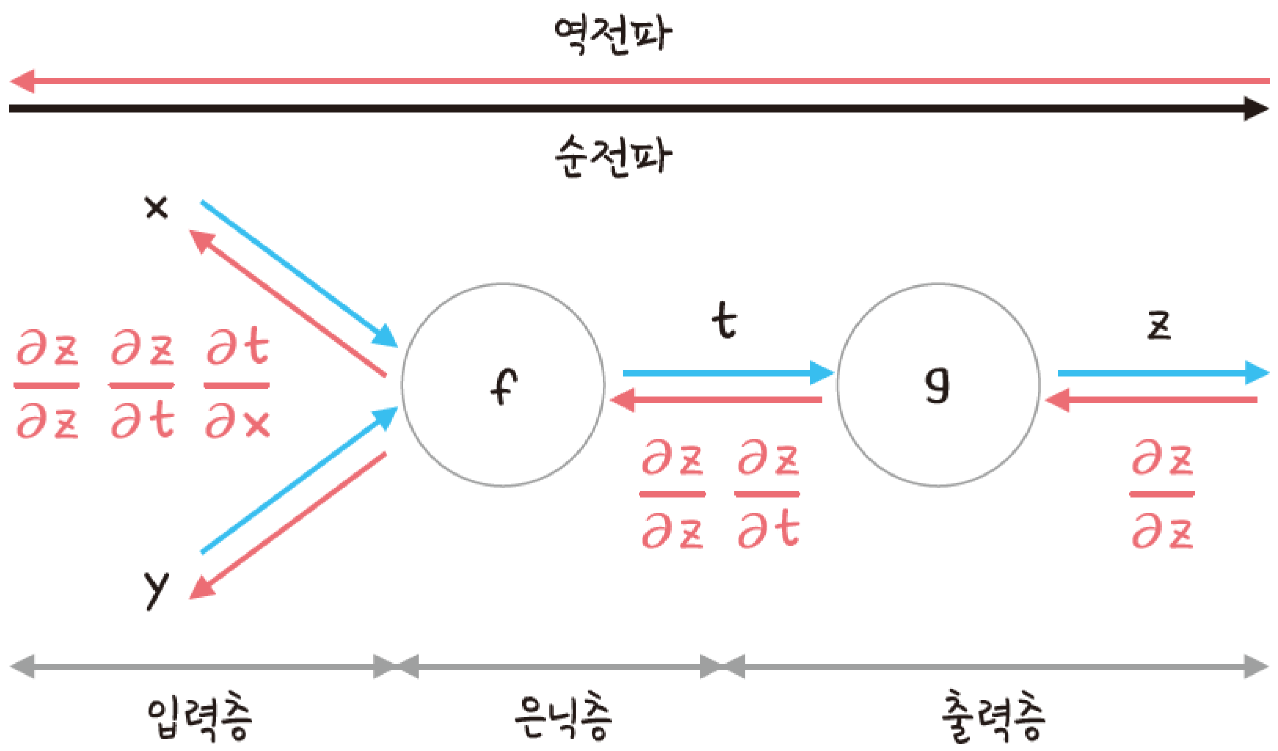
Differentiation and backpropagation

❖ Backpropagation

- The calculation method of error backpropagation is as follows
  - (1) The input value multiplied by the weight and the bias are combined, and if the value exceeds the threshold of 0, 1 is outputted, and otherwise 0 is outputted
  - (2) Obtain the error that is the difference between the output value and the correct answer
    - Correction to a weight value that reduces the error in the reverse direction (how the differentiation method expressed in red obtains the error)
  - (3) Modifies the weight value of the output layer
  - (4) Modify the weight value of the hidden layer

❖ Backpropagation

- Repeat steps (2) through (4) until the error is no longer reduced



Differentiation and backpropagation

❖ Backpropagation

- The backpropagation is said to correct the weight immediately before it in the direction of reducing the error
- When correcting the weight, the partial differential value of y = g(f(x) calculated from the net propagation is multiplied by the error and transmitted to the downstream node (the hidden layer)
- The reason for using **partial differential** at this time is that it does not need to consider all the weight values given to numerous nodes, only the connected weights need to be considered
- This is because using the chain law, even if there are many hidden layers between the output layer and the input layer, the slope can be calculated with a simple differential

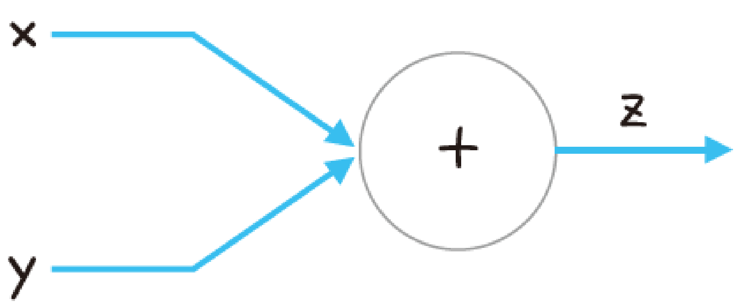
Differentiation and backpropagation

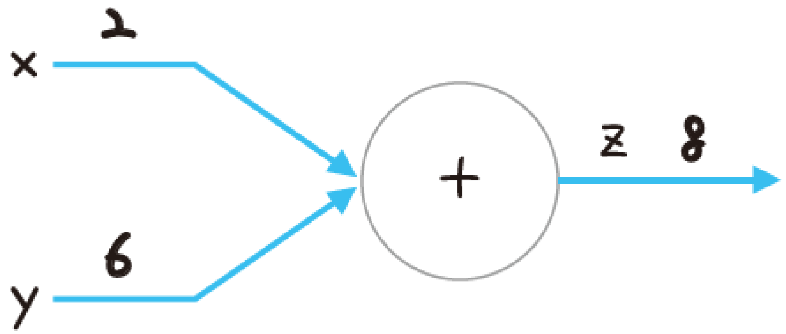
❖ Backpropagation

Backpropagation calculation

Node backpropagation of addition

- The calculation graph for the expression z = x + y is as follows
- The forward propagation calculation for z = x + y when x = 2, y = 6 is as follows





Differentiation and backpropagation

❖ Backpropagation

- It's not difficult because you can calculate the forward propagation sequentially
- Backpropagation calculations can be difficult to differentiate
- The derivative of the backpropagation of the addition is

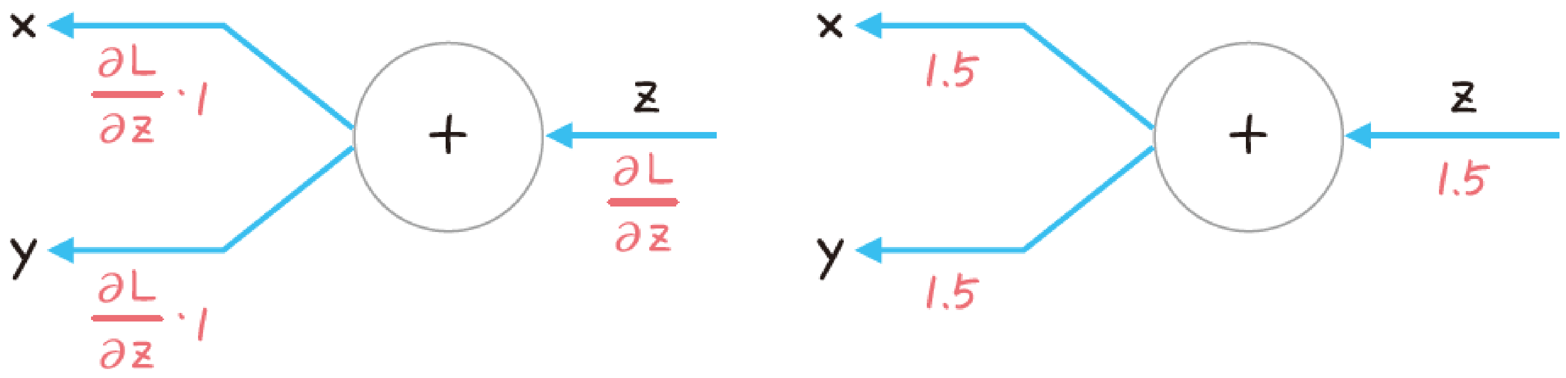
$$z = x + y$$

$$\frac{\partial z}{\partial x} = 1, \frac{\partial z}{\partial y} = 1$$

Differentiation and backpropagation

❖ Backpropagation

- Differentiable (L is considered the final output value)



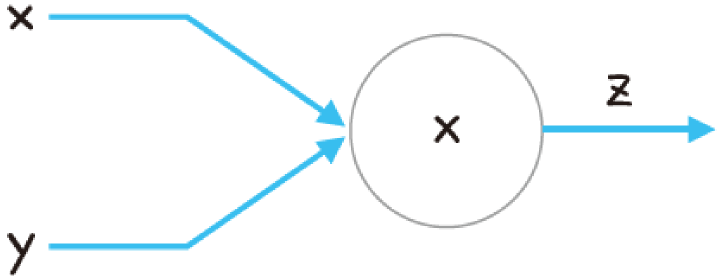
- Assuming that a value of 1.5 is input at upstream (output), the backpropagation of the addition node propagates the input value to the next node

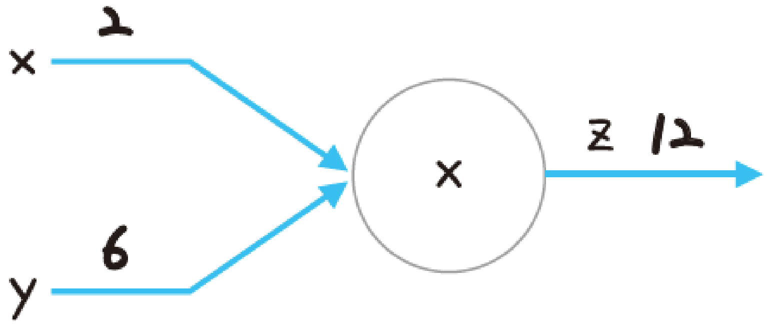
Differentiation and backpropagation

❖ Backpropagation

Node backpropagation of multiplication

- The calculation graph for z = xy expression is shown on the left
- The forward propagation calculation for z = xy when x = 2, y = 6 is equal to the right





Differentiation and backpropagation

❖ Backpropagation

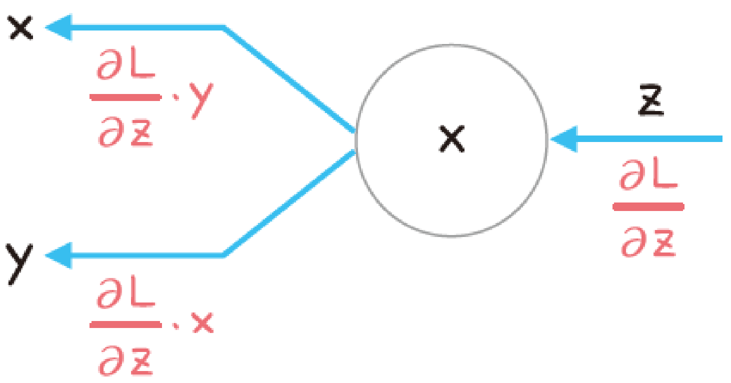
- The forward propagation of multiplication is also not difficult to calculate sequentially, but backpropagation requires knowing how to differentiate
- The derivative of the backpropagation of multiplication is

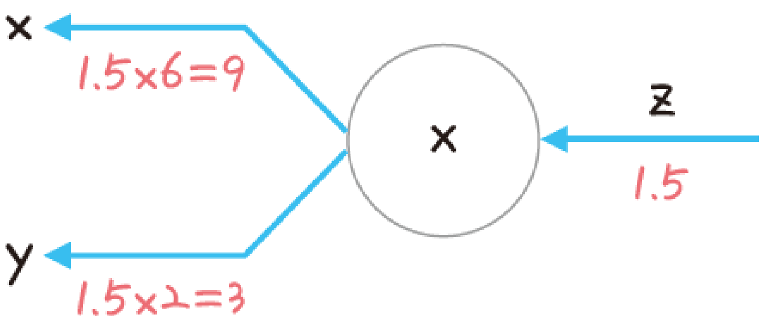
$$z = xy$$

$$\frac{\partial z}{\partial x} = y, \ \frac{\partial z}{\partial y} = x$$

❖ Backpropagation

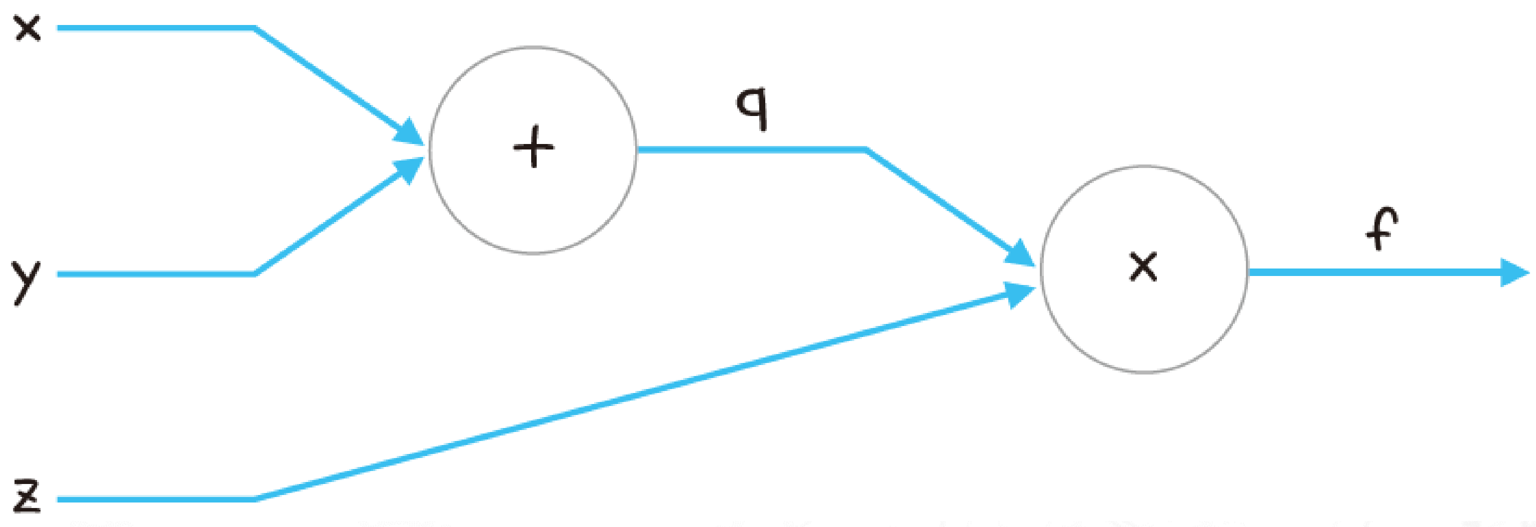
- Differential is possible as shown on the left
- Assuming that a value of 1.5 is input from the upstream, the backpropagation of the multiplication node can be sent downstream (the hidden layer) by multiplying the upstream (output) value by the input signal of the forward propagation by the 'switched value'





❖ Backpropagation

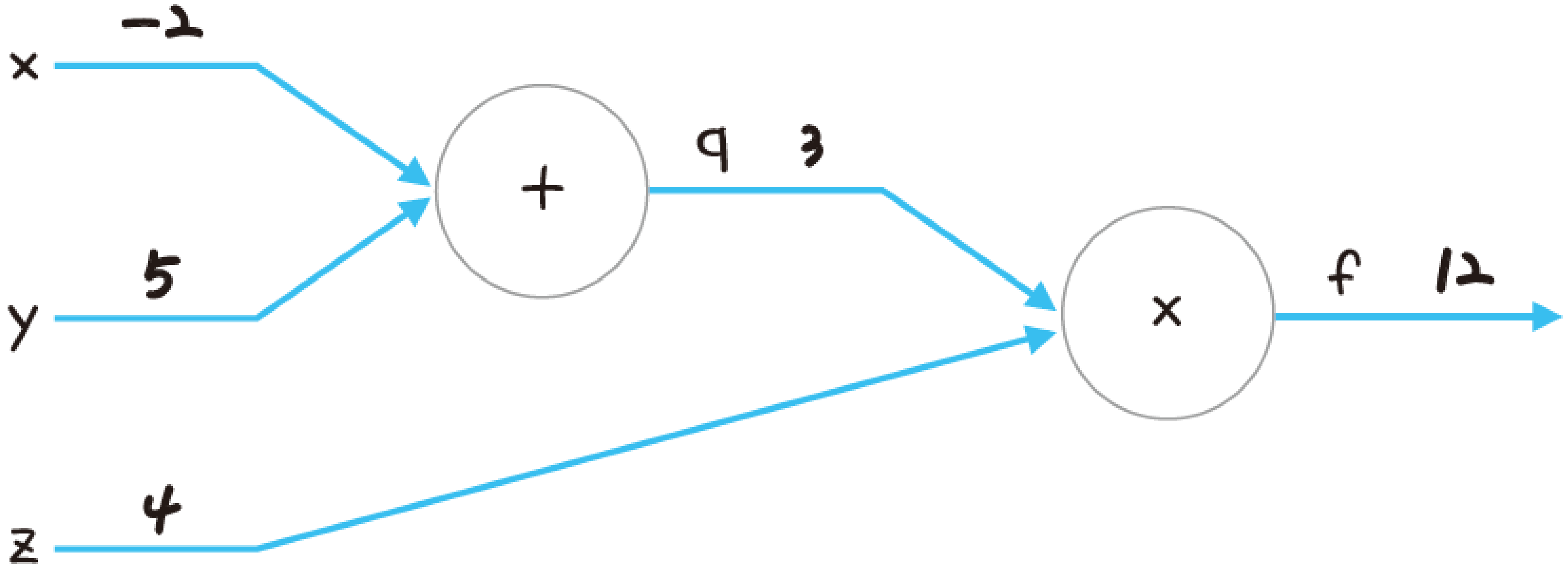
- Let's combine the backpropagation of addition and multiplication and look at it as a concrete example
- For example, if you have two functions, $q(x) = x + y$, $f(x) = q(x) \times z$, and you graph them with computational graphs



Differentiation and backpropagation

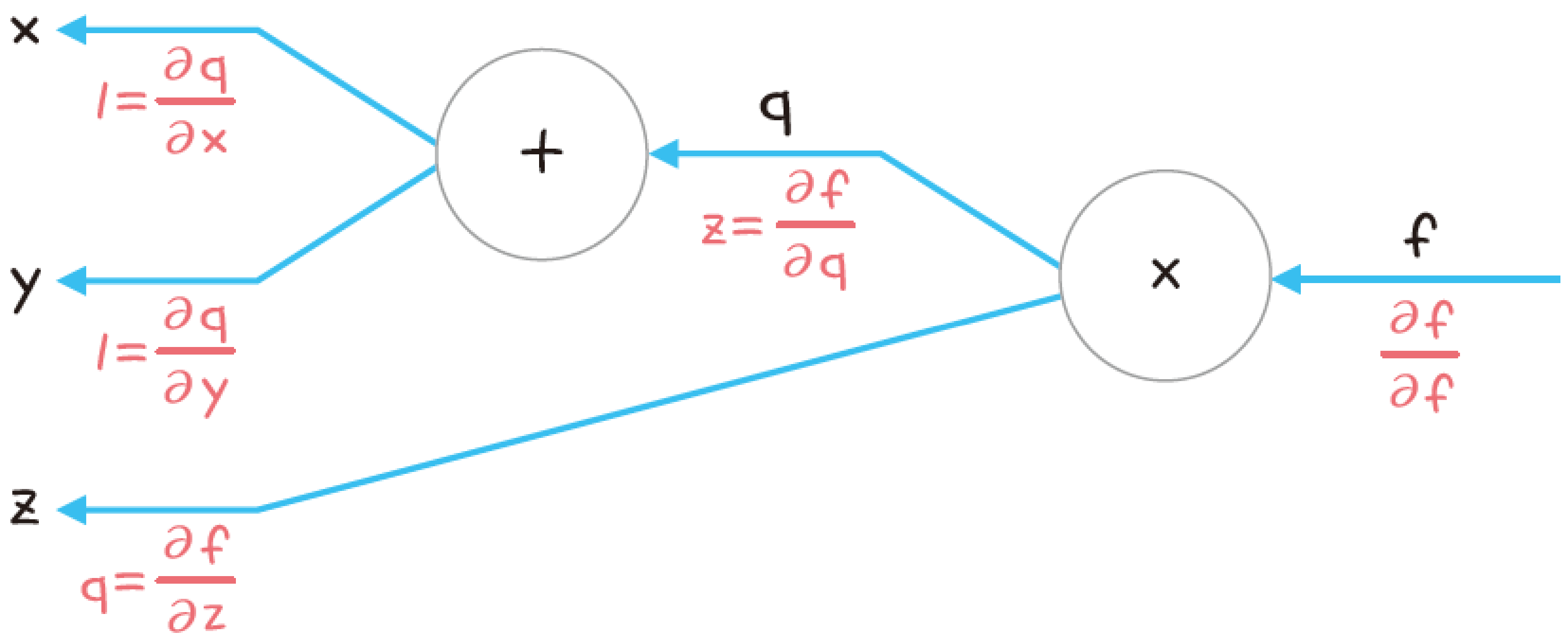
❖ Backpropagation

- If the forward calculation is performed when x = -2, y = 5, z = 4, the result is 12 as follows



Backpropagation

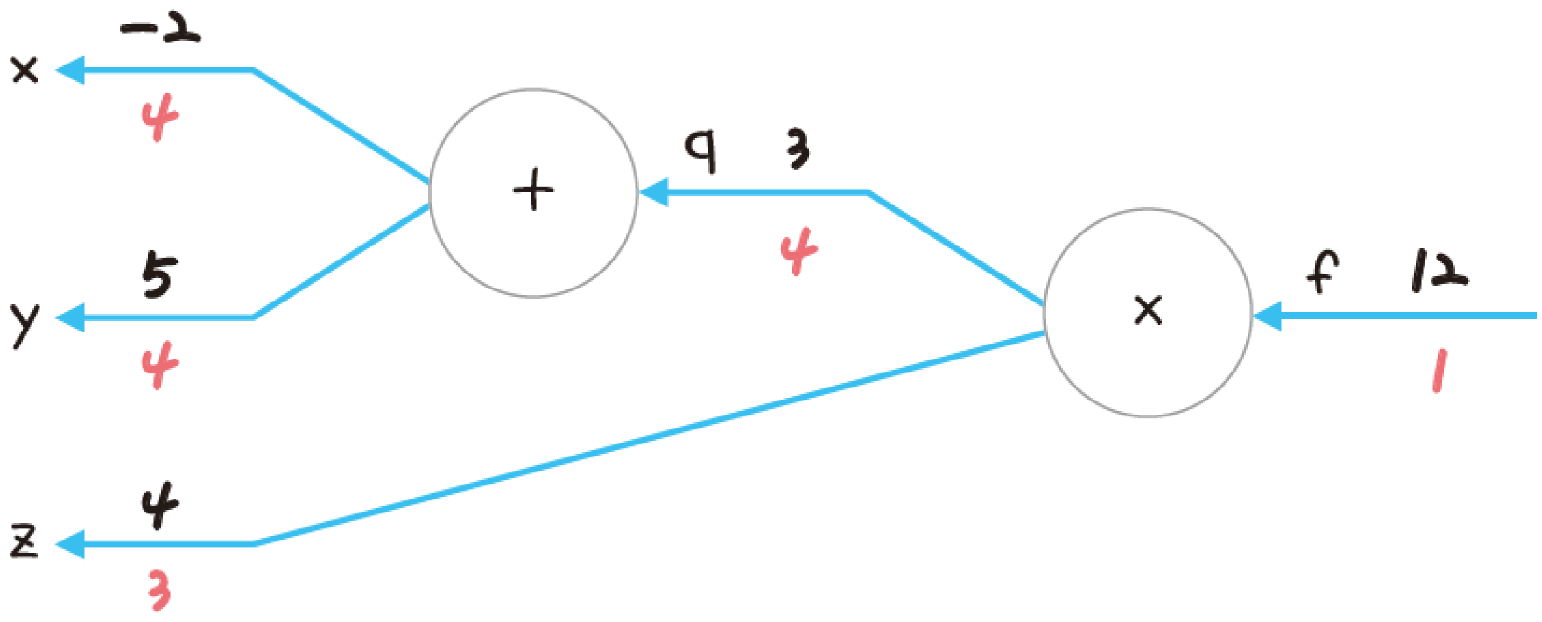
- If differential is applied to calculate the backpropagation, it can be differentiated as follows



Differentiation and backpropagation

❖ Backpropagation

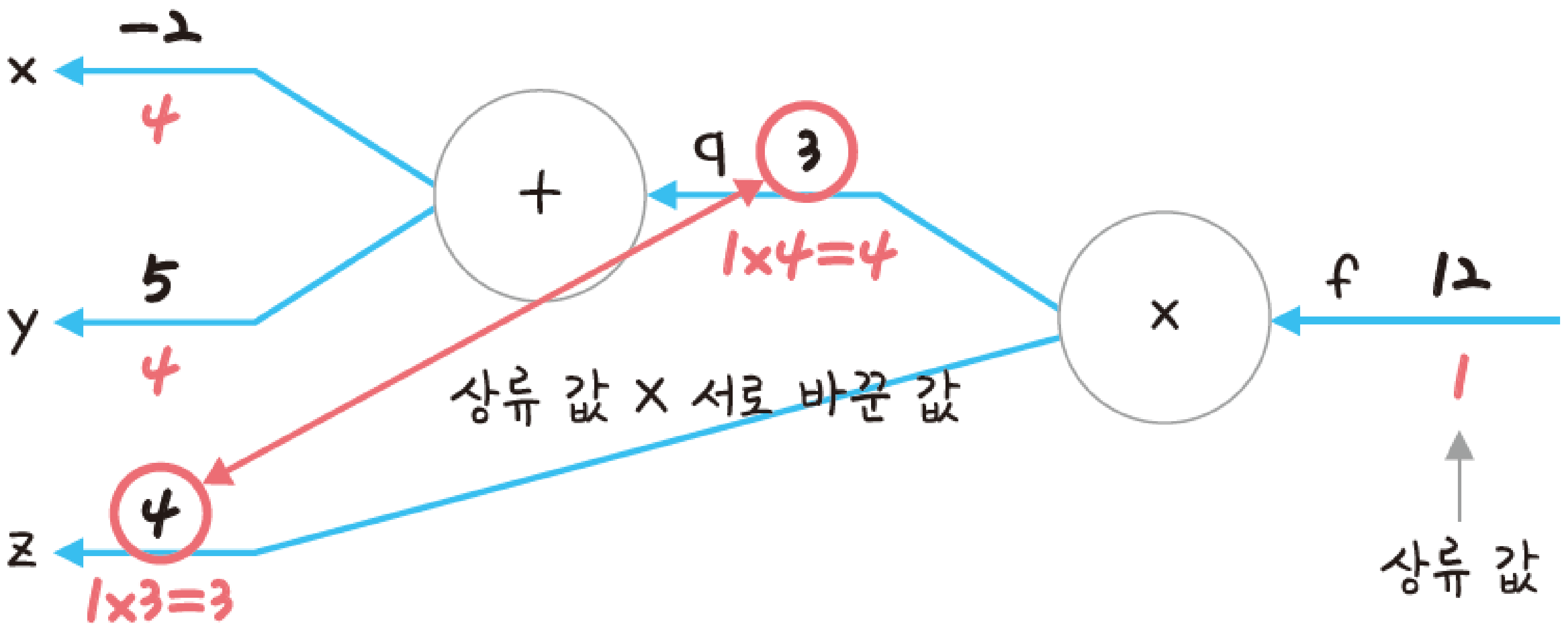
- In other words, for backpropagation of multiplication, the upstream (output) value (starting from 1) is multiplied by the input signal of the net propagation by the 'switched value' and sent downstream (the hidden layer)
- If you send the input value downstream (input) for backpropagation of the addition, the following results are obtained



Differentiation and backpropagation

❖ Backpropagation

- If you look at the previous results in detail, first, the backpropagation for multiplication is as follows
- You can replace the q-value with the z-value and multiply it with the upstream value



❖ Backpropagation

- The backpropagation for addition is