# Lecture A1. Math Review

Sim, Min Kyu, Ph.D., mksim@seoultech.ac.kr

서울과학기술대학교 데이터사이언스학과

# I. Differentiation and Integration

# Differentiation

**Definition 1 (differentiation)**

Differentiation is the action of computing a derivative.

**Definition 2 (derivative)**

The derivative of a function $y = f(x)$ of a variable $x$ is a measure of the rate at which the value $y$ of the function changes with respect to (wrt., hereafter) the change of the variable $x$. It is notated as $f'(x)$ and called derivative of $f$ wrt. $x$.

**Remark 1**

If $x$ and $y$ are real numbers, and if the graph of $f$ is plotted against $x$, the derivative is the slope of this graph at each point.

### Definition 3 (differentiable)

If $\lim_{h\to 0} \frac{f(x+h/2)-f(x-h/2)}{h}$ exists for a function $f$ at $x$, we say the function $f$ *is differentiable at* $x$. That is, $f'(x) = \lim_{h\to 0} \frac{f(x+h/2)-f(x-h/2)}{h}$. If $f$ is differentiable for all $x$, then we say $f$ *is differentiable (everywhere)*.

### Remark 2

The followings are popular derivatives.

- $f(x) = x^p \Rightarrow f'(x) = px^{p-1}$ (polyomial)
- $f(x) = e^x \Rightarrow f'(x) = e^x$ (exponential)
- $f(x) = log(x) \Rightarrow f'(x) = 1/x$ (log function; not differentiable at $x = 0$)

### Theorem 1

*Differentiation is linear. That is,* $h(x) = f(x) + g(x)$ *implies* $h'(x) = f'(x) + g'(x)$.

**Theorem 2 (differentiation of product)**

*If* $h(x) = f(x)g(x)$, *then* $h'(x) = f'(x)g(x) + f(x)g'(x)$.

**Exercise 1**

*Suppose* $f(x) = xe^x$, *find* $f'(x)$.

$f(x) = xe^x$
$f'(x) = e^x + xe^x$

**Theorem 3 (differentiation of fraction)**

*If* $h(x) = \frac{f(x)}{g(x)}$, *then* $h'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{(g(x))^2}$.

### Theorem 4 (composite function)

*If* $h(x) = f(g(x))$, *then* $h'(x) = f'(g(x)) \cdot g'(x)$.

### Exercise 2

*Suppose* $f(x) = e^{2x}$, *find* $f'(x)$.

$g(x) = e^x \quad f(x) = g(2x)$

$f'(x) = g'(2x) \times 2 = 2e^{2x}$

# Integration

**Definition 4 (integration)**

Integration is the computation of an integral, which is a reverse operation of differentiation up to an additive constant.

**Definition 5 (antiderivative)**

Let's say a function $f$ is a derivative of $g$, or $g'(x) = f(x)$, then we say $g$ is an *antiderivative* of $f$, written as $g(x) = \int f(x)dx + C$, where $C$ is a integration constant.

### Remark 3

The followings are popular antiderivatives.

- For $p \neq 1$, $f(x) = x^p \Rightarrow \int f(x)dx = \frac{1}{p+1}x^{p+1} + C$ (polyomial)
- $f(x) = \frac{1}{x} \Rightarrow \int f(x)dx = log(x) + C$ (fraction)
- $f(x) = e^x \Rightarrow \int f(x)dx = e^x + C$ (exponential)
- $f(x) = \frac{g'(x)}{g(x)} \Rightarrow \int f(x)dx = log(g(x)) + C$ (See Theorem 4 above)

### Exercise 3

*Derive $\int f'(x)g(x)\ dx = f(x)g(x) - \int f(x)g(x)'\ dx$. (Hint: Use Theorem 2 above.)*

$\int f'(x)g(x) = f(x)g(x) - \int f(x)g'(x)$

$f(x)g(x) = \int f'(x)g(x) + \int f(x)g'(x)$ ← differentiate

$f'(x)g(x) + f(x)g'(x) = f'(x)g(x) + f(x)g'(x)$

## Exercise 4

*Find* $\int xe^x \; dx$*, and evaluate* $\int_0^1 xe^x \; dx$*. (Hint: Use Exercise 3 above.)*

$$\int f'(x)g(x) = f(x)g(x) - \int f(x)g'(x)$$

$$\int f(x)g'(x) = f(x)g(x) - \int f'(x)g(x)$$

$f(x) = x \quad g'(x) = e^x$

$f'(x) = 1 \quad g(x) = e^x + c$

$$= x(e^x + c) - \int e^x + c$$

$$= x(e^x + c) - (e^x + cx + D)$$

$$= xe^x + \cancel{cx} - e^x - \cancel{cx} - D$$

$$= xe^x - e^x + C$$

$$\int_0^1 xe^x = \left[ xe^x - x + C \right]_0^1$$

$$= (e - 1 + C) - C$$

$$= e - 1$$

# II. Numerical Methods

## Differentiation

- Oftentimes, finding analytic derivative is hard, but finding numerical derivative is often possible.

### Definition 6

For a function $f$ and small constant $h$,

- $f'(x) \approx \frac{f(x+h)-f(x)}{h}$ *(forward difference formula)*
- $f'(x) \approx \frac{f(x)-f(x-h)}{h}$ *(backward difference formula)*
- $f'(x) \approx \frac{f(x+h)-f(x-h)}{2h}$ *(centered difference formula)*

# Solving an equation

- For the rest of this section, we consider a nonlinear and differentiable (thus, continuous) function $f : \mathbb{R} \to \mathbb{R}$, we aim to find a point $x^* \in \mathbb{R}$ such that $f(x^*) = 0$. We call such $x^*$ as a *solution* or a *root*.

# Bisection Method

- The *bisection* method aims to find a very short interval $[a, b]$ in which $f$ changes a sign.
- Why? Changing a sign from $a$ to $b$ means the function crosses the $\{y = 0\}$-axis, (a.k.a. $x$-axis), at least once. It means $x^*$ such that $f(x^*) = 0$ is in this interval. Since $[a, b]$ is a very short interval, We may simply say $x^* = \frac{a+b}{2}$.

**Definition 7 (sign function)**

$sgn(\cdot)$ is called a *sign function* that returns 1 if the input is positive, -1 if negative, and 0 if zero.

# Bisection algorithm

- Let $tol$ be the maximum allowable length of the *short interval* and an initial interval $[a, b]$ be such that $sgn(f(a)) \neq sgn(f(b))$.
- The *bisection algorithm* is the following.

```
while ((b − a) > tol) do
    m = (a+b)/2
    if sgn(f(a)) = sgn(f(m)) then
        a = m
    else
        b = m
    end
end
```

- At each *iteration*, the interval length is halved. As soon as the interval length becomes smaller than $tol$, then the algorithm stops.

# Newton Method

- The bisection technique makes no used of the function values other than their signs, resulting in slow but sure convergence.
- More rapid convergence can be achieved by using the function values to obtain a more accurate approximation to the solution *at each iteration*.
- Newton method is a method that use both the function value and derivative value.

- Newton method approximates the function $f$ near $x_k$ by the tangent line at $f(x_k)$.

1: $x_0$ = initial guess
2: for $k$=0,1,2,...
3: $\quad x_{k+1} = x_k - f(x_k)/f'(x_k)$
4: $\quad$ break if $|x_{k+1} - x_k| < tol$
5: end

- Root-finding numerical methods such as bisection method and newton method has a few common properties.
  1. It is characterized as a *iterative process* (such as $x_0 \to x_1 \to x_2 \to \cdots$).
  2. In each *iteration*, the current candidate *gets closer* to the true value.
  3. It converges. That is, it is theoretically reach the *exact value* up to tolerance.

- Many iterative numerical methods share the properties above.
- The famous back propagation in deep neural network is also motivated by Newton method.
- Major algorithms for dynamic programming are called *policy iteration* and *value iteration* that also share the properties above.

# III. Matrix Algebra

# Matrix multiplication

### Exercise 5

*Solve the followings.*

$$(.6 \quad .4)\begin{pmatrix} .7 & .3 \\ .5 & .5 \end{pmatrix} =$$

$(.42 + .2, \ .18 + .20)$

$= (.62, \ .38)$

## Exercise 6

*What is $P^2$?*

$$P = \begin{pmatrix} .7 & .3 \\ .5 & .5 \end{pmatrix}$$

$$\begin{pmatrix} .7 & .3 \\ .5 & .5 \end{pmatrix}\begin{pmatrix} .7 & .3 \\ .5 & .5 \end{pmatrix} = \begin{pmatrix} .49 + .15, & .21 + .15 \\ .35 + .25, & .15 + .25 \end{pmatrix}$$

$$= \begin{pmatrix} .64 & .36 \\ .6 & .4 \end{pmatrix}$$

# Solution to system of linear equations

**Exercise 7**

*Solve the followings.*

$$\begin{pmatrix}\pi_1 & \pi_2\end{pmatrix}\begin{pmatrix}.7 & .3\\ .5 & .5\end{pmatrix} = \begin{pmatrix}\pi_1 & \pi_2\end{pmatrix}$$

$$\pi_1 + \pi_2 = 1$$

$= (.7\pi_1 + .5\pi_2 \,,\ .3\pi_1 + .5\pi_2) = (\pi_1, \pi_2)$

$.7\pi_1 + .5\pi_2 = \pi_1 \qquad .3\pi_1 + .5\pi_2 = \pi_2$

$.5\pi_2 = .3\pi_1 \qquad .3\pi_1 = .5\pi_2 \,, \quad \pi_1 = \frac{5}{3}\pi_2$

$\pi_1 + \pi_2 = 1 \qquad \frac{8}{3}\pi_2 = 1 \qquad \boxed{\pi_2 = \frac{3}{8},\ \pi_1 = \frac{5}{8}}$ //

## Exercise 8 ?

*Solve the following system of equations.*

$$x = y$$
$$y = 0.5z \quad (-)?$$
$$z = 0.6 + 0.4x$$
$$x + y + z = 1$$

$x=y$ , $y=0.5z$

$2y+z=1$

$2z=1$

$z=1/2$ $y=1/4$ $x=1/4$

$z=0.6+0.4x$

$z=0.7$

$z=0.5?$

$\frac{4}{10}\cdot\frac{1}{4}=0.1$

## Exercise 9

*Solve the following system of equations.*

$$\begin{pmatrix}\pi_0 & \pi_1 & \pi_2\end{pmatrix}\begin{pmatrix}-2 & 2 & 0\\ 3 & -5 & 2\\ 0 & 3 & -3\end{pmatrix} = \begin{pmatrix}0 & 0 & 0\end{pmatrix}$$

$$\pi_0 + \pi_1 + \pi_2 = 1$$

$$-2\pi_0 + 3\pi_1 = 0$$

$$2\pi_0 - 5\pi_1 + 3\pi_2 = 0$$

$$2\pi_1 - 3\pi_2 = 0$$

$3\pi_1 = 2\pi_0 \quad \pi_1 = \frac{2}{3}\pi_0$

$3\pi_2 = 2\pi_1 \quad \pi_2 = \frac{2}{3}\pi_1 = \frac{4}{9}\pi_0$

$\frac{18}{19} - \frac{30}{19} + \frac{12}{19} = 0$ ✓

$$\pi_0 + \pi_1 + \pi_2 = 1$$

$$\pi_0 + \frac{2}{3}\pi_0 + \frac{4}{9}\pi_0 = 1$$

$$\frac{19}{9}\pi_0 = 1$$

$$\pi_0 = \frac{9}{19} \quad \pi_1 = \frac{6}{19} \quad \pi_2 = \frac{4}{19}$$

## Exercise 10

*Solve the following system of equations.*

$$\begin{pmatrix}\pi_1 & \pi_2 & \pi_3 & \pi_4\end{pmatrix}\begin{pmatrix} .7 & .3 & & \\ .5 & .5 & & \\ & & .6 & .4 \\ & & .3 & .7 \end{pmatrix} = \begin{pmatrix}\pi_1 & \pi_2 & \pi_3 & \pi_4\end{pmatrix}$$

$$\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$$

$.7\pi_1 + .5\pi_2 = \pi_1$

$.3\pi_1 + .5\pi_2 = \pi_2$

$.4\pi_1 = \pi_1 - \pi_2$

$\pi_2 = .6\pi_1$

$.6\pi_3 + .3\pi_4 = \pi_3$

$.4\pi_3 + .7\pi_4 = \pi_4$

$0.4\pi_3 = 0.3\pi_4$

$.4\pi_3 = .3\pi_4$

$\pi_3 = 3/4\,\pi_4$

$8/5\,\pi_2 + 7/4\,\pi_4 = 1$

## Exercise 11

*Solve following and express $\pi_i$ for $i = 0, 1, 2, ...$*

$$\begin{aligned}
\pi_0 + \pi_1 + \pi_2 + ... &= 1 \\
0.02\pi_0 + 0.02\pi_1 + 0.02\pi_2 + ... &= \pi_0 \\
0.98\pi_0 &= \pi_1 \\
0.98\pi_1 &= \pi_2 \\
0.98\pi_2 &= \pi_3 \\
... &= ...
\end{aligned}$$

$\pi_1 = 0.98\,\pi_0$

$\pi_2 = 0.98\,\pi_1 = (0.98)^2\,\pi_0$

$\pi_3 = (0.98)^3\,\pi_0$

$\vdots$

$\pi_i = (0.98)^i\,\pi_0$

$\sum_{i=0}^{n} \pi_i = \frac{\pi_0(1-r^n)}{1-r} = 1$

$= \frac{\pi_0(1-0.98^n)}{0.02}$

$1 = \frac{\pi_0}{0.02}$

$\pi_0 = 0.02$

$\therefore\ \pi_i = (0.02)(0.98)^i$

# IV. Series and Others

## Exercise 12 (Infinite geometric series)

*Simplify the following. When* $|r| < 1$, $S = a + ar + ar^2 + ar^3 + \ldots$

$S = a + ar + ar^2 + \cdots$

$rS = ar + ar^2 \cdots$

$ar^{n+1} \to 0$

$|r| < 1$

$(1-r)S = a$

$S = \frac{a}{1-r}$

$\boxed{S = \frac{a}{1-r}}$

## Exercise 13 (Finite geometric series)

*Simplify the following. When* $r \neq 1$, $S = a + ar + ar^2 + ar^3 + ... + ar^{n-1}$

$$S = a + ar + ar^2 + \cdots ar^{n-1}$$
$$rS = ar + ar^2 + \cdots ar^{n-1} + ar^n$$

$$(1-r)S = a - ar^n$$

$$\boxed{S = \frac{a(1-r)^n}{1-r}}$$

## Exercise 14 (Power series)

*Simplify the following. When* $|r| < 1$, $S = r + 2r^2 + 3r^3 + 4r^4 + \ldots$

$$S = r + 2r^2 + 3r^3 + \cdots$$
$$rS = \quad r^2 + 2r^3 + \cdots$$

$$(1-r)S = r + r^2 + r^3 + \cdots$$

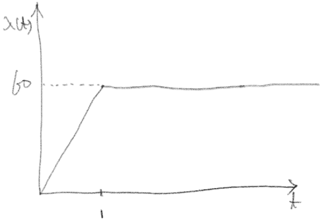
$$(1-r)S = \frac{1}{1-r}$$

$$\boxed{S = \frac{1}{(1-r)^2}}$$

# Formulation of time varying function

## Exercise 15

*During the first hour* $(0 \le t \le 1)$, $\lambda(t)$ *increases linearly from* $0$ *to* $60$. *After the first hour,* $\lambda(t)$ *is constant at* $60$. *Draw plot for* $\lambda(t)$ *and express the function in math form.*




$$\lambda(t) = \begin{cases} 60t & (0 \le t \le 1) \\ 60 & (t > 1) \end{cases}$$

```
"Man can learn nothing unless he proceeds from the known to the unknown. - Claude Bernard"
```