Lecture A1. Math Review

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

SEOULTECH 서울과학기술대학교 데이터사이언스학과

I. Differentiation and Integration

Differentiation

Definition 1 (differentiation) 미분

Differentiation is the action of computing a derivative.

Definition 2 (derivative) 적분

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) 미분가능

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) 다항식
- $f(x) = e^x \Rightarrow f'(x) = e^x$ (exponential) 지수함수
- $f(x) = log(x) \Rightarrow f'(x) = 1/x$ (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* $f'(x)$.

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

$$f'(x) = 1 \cdot e^x + xe^x$$

$$f'(x) = (1+x)e^x$$

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* $f'(x)$.

$$\left(e^{f(x)}\right)' = e^{f(x)} \cdot f'(x)$$

$$\left(e^{2x}\right)' = \left(e^{2x}\right)' \cdot (2x)'$$

$$= 2e^{2x}$$

Integration 적분

Definition 4 (integration)

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

미분

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.)*

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

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

$\int (f'(x)g(x))dx + \int (f(x)g'(x))dx = f(x)g(x)$

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

Exercise 4

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

Exercise 3 $\int f'(x)g(x)\,dx = f(x)g(x) - \int (f(x)g'(x))\,dx$

$$\int xe^x\,dx = e^x x - \int (e^x \cdot 1)\,dx$$
$$= e^x x - e^x + C$$

$$\int_0^1 xe^x\,dx = \left[e^x x - e^x - C\right]_0^1$$
$$= (e^1 \cdot 1 - e^1 - C) - (e^0 \cdot 0 - e^0 - C)$$
$$= (0 - C) - (0 - 1 - C)$$
$$= 1$$

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 $f(x_k)$.

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

$$(.6 \quad .4)\begin{pmatrix} .7 & .3 \\ .5 & .5 \end{pmatrix} = (.64 \quad .38)$$

$42 + 20 \qquad 18 + 20$

Exercise 6

What is P^2?

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

$$P^2 = \begin{pmatrix} 49+15 & 21+15 \\ 35+25 & 15+25 \end{pmatrix}$$

$$= \begin{pmatrix} .64 & .36 \\ .60 & .40 \end{pmatrix}$$

Solution to system of linear equations

Exercise 7

Solve the followings.

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

$$\pi_1 + \pi_2 = 1$$

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

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

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

$\Rightarrow$ $7\pi_1 + 5\pi_2 = 10\pi_1$

$3\pi_1 + 5\pi_2 = 10\pi_2$

① $-3\pi_1 + 5\pi_2 = 0$

$\Rightarrow$ ② $3\pi_1 - 5\pi_2 = 0$

$\Rightarrow$ $\underline{\pi_1 = \frac{5}{3}\pi_2}$ $\Rightarrow$ $\pi_2 = \frac{3}{8}$, $\pi_1 = \frac{5}{8}$

Exercise 8

Solve the following system of equations.

$$
\begin{aligned}
x &= y \\
y &= 0.5z \\
z &= 0.6 + 0.4x \\
x + y + z &= 1
\end{aligned}
$$

-로 수정 아 해가 없음

$$
\begin{aligned}
x - y &= 0 \\
10y - 5z &= 0 \\
10z - 4x &= 0
\end{aligned}
\Rightarrow
\begin{pmatrix} 1 & -1 & 0 \\ 0 & 10 & -5 \\ -4 & 0 & 10 \end{pmatrix}
\begin{pmatrix} x \\ y \\ z \end{pmatrix}
=
\begin{pmatrix} 0 \\ 0 \\ 6 \end{pmatrix}
\Rightarrow
\left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 10 & -5 & 0 \\ -4 & 0 & 10 & 6 \end{array}\right)
$$

$$
\Rightarrow
\left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 10 & -5 & 0 \\ 0 & -4 & 10 & 6 \end{array}\right)
\Rightarrow
\left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & -2 & 5 & 3 \end{array}\right)
\Rightarrow
\left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 4 & 3 \end{array}\right)
\Rightarrow
\left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 4 & 3 \end{array}\right)
\Rightarrow
\left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & \frac{3}{4} \end{array}\right)
$$

$$
\Rightarrow
\left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{3}{8} \\ 0 & 1 & 0 & \frac{3}{8} \\ 0 & 0 & 1 & \frac{3}{4} \end{array}\right)
\Rightarrow x = \frac{3}{8},\ y = \frac{3}{8},\ z = \frac{3}{4},
$$

$x+y+1=1$의 해는 없는 것 같습니다

Exercise 9

Solve the following system of equations.

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

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

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

$$\Rightarrow \left(\begin{array}{ccc|c} -2 & 3 & 0 & 0 \\ 2 & -5 & 3 & 0 \\ 0 & 2 & -3 & 0 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|c} -2 & 3 & 0 & 0 \\ 0 & -2 & 3 & 0 \\ 0 & 2 & -3 & 0 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|c} -2 & 3 & 0 & 0 \\ 0 & -2 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \Rightarrow \left(\begin{array}{ccc|c} 1 & -\frac{3}{2} & 0 & 0 \\ 0 & 1 & -\frac{3}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\begin{array}{l} \pi_0 - \frac{3}{2}\pi_1 = 0 \\ \pi_1 - \frac{3}{2}\pi_2 = 0 \end{array} \Rightarrow \begin{array}{l} \pi_0 = \frac{3}{2}\pi_1 \\ \pi_2 = \frac{2}{3}\pi_1 \end{array} \Rightarrow \frac{3}{2}\pi_1 + \pi_1 + \frac{2}{3}\pi_1 = 1 \Rightarrow \boxed{\pi_1 = \frac{6}{19},\ \pi_0 = \frac{9}{19},\ \pi_2 = \frac{4}{19}}$$

Exercise 10

Solve the following system of equations.

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

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

$$\frac{5}{8} - \frac{35}{32}\pi_4 + \frac{3}{8} - \frac{21}{32}\pi_4 + \frac{3}{4}\pi_4 + \pi_4 = 1$$

$$\frac{\cancel{8}}{\cancel{8}} - \frac{36}{32}\pi_4 + \frac{7}{4}\pi_4 = \cancel{1}$$

$$-36\pi_4 + 56\pi_4 = 0$$

$$20\pi_4 = 0 \ ? \quad \pi_4 = 0$$

$$\therefore \ \pi_1 = \frac{5}{8},\ \pi_2 = \frac{3}{8},\ \pi_3 = 0,\ \pi_4 = 0$$

$$\Rightarrow \begin{pmatrix} .7 & .5 & 0 & 0 \\ .3 & .5 & 0 & 0 \\ 0 & 0 & .6 & .3 \\ 0 & 0 & .4 & .7 \end{pmatrix}\begin{pmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{pmatrix} = \begin{pmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \pi_4 \end{pmatrix}$$

$$\Rightarrow \begin{aligned} 0.7\pi_1 + 0.5\pi_2 &= \pi_1 \\ 0.3\pi_1 + 0.5\pi_2 &= \pi_2 \\ 0.6\pi_3 + 0.3\pi_4 &= \pi_3 \\ 0.4\pi_3 + 0.7\pi_4 &= \pi_4 \end{aligned} \quad \Rightarrow \begin{aligned} 3\pi_1 - 5\pi_2 &= 0 \\ -3\pi_1 + 5\pi_2 &= 0 \\ 4\pi_3 - 3\pi_4 &= 0 \\ 3\pi_4 - 4\pi_3 &= 0 \end{aligned} \quad \Rightarrow \begin{aligned} \pi_1 &= \frac{5}{3}\pi_2 \\ \pi_3 &= \frac{3}{4}\pi_4 \end{aligned}$$

$$\Rightarrow \frac{5}{3}\pi_2 + \pi_2 + \frac{3}{4}\pi_4 + \pi_4 = 1$$

$$\rightarrow \frac{8}{3}\pi_2 + \frac{7}{4}\pi_4 = 1$$

$$\Rightarrow \rightarrow 32\pi_2 + 21\pi_4 = 12$$

$$\begin{pmatrix} \pi_2 = \frac{3}{8} - \frac{21}{32}\pi_4 \\ \pi_1 = \frac{5}{8} - \frac{35}{32}\pi_4 \end{pmatrix}$$

Exercise 11

Solve following and express π_i for $i = 0, 1, 2, ...$

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

Pass ...

IV. Series and Others

Exercise 12 (Infinite geometric series)

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

$$S_n = a + \cancel{ar} + \cancel{ar^2} + \cdots + \cancel{ar^{n-1}} \qquad S = a \cdot r^{n-1}$$

$$- \quad rS_n = \cancel{ar} + \cancel{ar^2} \cdots + \cancel{ar^{n-1}} + ar^n$$

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

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

만약 n이 무한(infinite)이면

$$\therefore \sum_{n=1}^{\infty} S_n = \lim_{n \to \infty} S_n = \frac{a}{1-r}$$

Exercise 13 (Finite geometric series)

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

$$S_n = a + \cancel{ar} + \cancel{ar^2} + \cdots + \cancel{ar^{n-1}}$$

$$- \quad rS_n = \cancel{ar} + \cancel{ar^2} + \cdots + \cancel{ar^{n-1}} + ar^n$$

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

$$\therefore S_n = \frac{a(1-r^n)}{1-r}, \; r \neq 1$$

Exercise 14 (Power series)

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

$$r + r^2 + r^3 + r^4 + \cdots = \frac{r}{1-r}$$

$$r^2 + r^3 + r^4 + \cdots = \frac{r^2}{1-r}$$

$$r^3 + r^4 + \cdots = \frac{r^3}{1-r}$$

$$\vdots \qquad \vdots$$

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

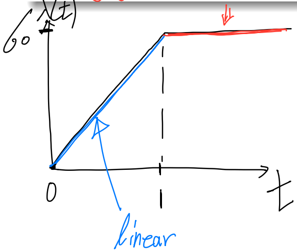
$$S = \frac{1}{1-r} \times \frac{r}{1-r} = \frac{r}{(1-r)^2}$$

Good!

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.*



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

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