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* $f'(x)$*.*

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

$= e^x + xe^x$

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* $f'(x)$.

$f(x) = e^{2x}$. $g(x) = e^{x}$, $h(x) = 2x$

$\Rightarrow f(x) = g(h(x))$

$f'(x) = g'(h(x)) \cdot h'(x)$

$= e^{2x} \cdot 2 \quad = 2e^{2x}$

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.)*

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

antiderivative

$$\Rightarrow \quad f(x)g(x) = \underline{\int f'(x)g(x)} + \int f(x)\cdot g'(x)$$

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

Exercise 4

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

$\int x e^x$ $\quad x = g(x)$, $e^x = h'(x)$

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

$\Rightarrow \int x \cdot e^x = x \cdot e^x - \int 1 \cdot e^x \cdot dx$

$\therefore \int x \cdot e^x = xe^x - e^x$

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

$= 0 - (-1) = 1$

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.
이분법
- 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
    $m = \frac{a+b}{2}$
    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: $x_{k+1} = x_k - f(x_k)/f'(x_k)$
4: 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.

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

$$\underbrace{(.6 \quad .4)}_{1\times 2}\underbrace{\begin{pmatrix} .7 & .3 \\ .5 & .5 \end{pmatrix}}_{2\times 2} = \underbrace{(.42+.20 \;,\; .18+.20)}_{1\times 2} = (.62 \quad .38)$$

Exercise 6

What is P^2?

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

$$P^2 = \begin{pmatrix} .7 & .3 \\ .5 & .5 \end{pmatrix}\begin{pmatrix} .7 & .3 \\ .5 & .5 \end{pmatrix} = \begin{pmatrix} .49+.15 & .21+.15 \\ .35+.25 & .15+.25 \end{pmatrix}$$

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

Solution to system of linear equations

Exercise 7

Solve the followings.

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

$$\pi_1 + \pi_2 = 1$$

$$(\pi_1 \quad \pi_2)\left(\begin{array}{c|c} .7 & .3 \\ .5 & .5 \end{array}\right) = (.7\pi_1 + .5\pi_2 \quad .3\pi_1 + .5\pi_2) = (\pi_1 \quad \pi_2)$$

$$\Rightarrow \quad .3\pi_1 - .5\pi_2 = 0 \quad \cdots ①$$

$$-.3\pi + .5\pi_2 = 0 \quad \cdots ②$$

$$\underline{\pi_1 = 1 - \pi_2} \qquad ① \Rightarrow .3(1-\pi_2) - .5\pi_2 = .3 - .8\pi_2 = 0$$

$$\pi_2 = \frac{.3}{.8} \text{ and } \pi_1 = \frac{.5}{.8}$$

Exercise 8

Solve the following system of equations.

$$x = y \quad \cdots \ 1$$
$$y = 0.5z \quad \cdots \ 2$$
$$z = 0.6 + 0.4x \quad \cdots \ 3$$
$$x + y + z = 1 \quad \cdots \ 4$$

$2 \to 1 \qquad x = 0.5z \quad \cdots \ 5$

$5 \to 3 \qquad z = 0.6 + 0.4 \cdot 0.5z, \quad z = 0.6 + 0.2z \qquad z = \frac{0.6}{0.8} = \frac{3}{4}$

$y = \frac{3}{8}, \quad x = \frac{3}{8}.$

but. $\frac{3}{8} + \frac{3}{8} + \frac{6}{8} = \frac{12}{8}$. this pair does not satisfy the equation 4

$\Rightarrow$ impossible

Exercise 9

Solve the following system of equations.

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

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

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

$-2\pi_0 + 3\pi_1 = 0$ ... ①

$2\pi_0 - 5\pi_1 + 3\pi_2 = 0$ ... ②

$2\pi_1 - 3\pi_2 = 0$ ... ③

$\pi_0 + \pi_1 + \pi_2 = 1$ ... ④

① $\Rightarrow \pi_0 = \frac{3}{2}\pi_1$ ③ $\Rightarrow \pi_2 = \frac{2}{3}\pi_1$

④ $\Rightarrow \frac{3}{2}\pi_1 + \pi_1 + \frac{2}{3}\pi_1 = 1$ $\therefore \underline{\pi_1 = \frac{6}{19}}$

$\underline{\pi_0 = \frac{9}{19}}$ $\underline{\pi_2 = \frac{4}{19}}$

② $\Rightarrow \frac{18}{19} - \frac{30}{19} + \frac{12}{19} = 0$

Exercise 10

Solve the following system of equations.

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

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

$$\left.\begin{array}{l} .7\pi_1 + .5\pi_2 = \pi_1 \Rightarrow .3\pi_1 - .5\pi_2 = 0 \\ .3\pi_1 + .5\pi_2 = \pi_2 \Rightarrow .3\pi_1 - .5\pi_2 = 0 \end{array}\right\} \text{same equation}$$

$$\left.\begin{array}{l} .6\pi_3 + .3\pi_4 = \pi_3 \Rightarrow .4\pi_3 - .3\pi_4 = 0 \\ .4\pi_3 + .7\pi_4 = \pi_4 \Rightarrow .4\pi_3 - .3\pi_4 = 0 \end{array}\right\} \text{same equation}$$

$\Rightarrow$ 4 variable, 3 equation : indeterminate.

Exercise 11

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

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

$$\begin{aligned}
0.02\pi_0 + 0.02\pi_1 + 0.02\pi_2 + \cdots &= \pi_0 \\
0.98\pi_0 &= \pi_1 \\
0.98\pi_1 &= \pi_2 \\
&\vdots \\
+ \quad & \\
\hline
\pi_0 + \pi_1 + \pi_2 + \pi_3 \cdots &= \pi_0 + \pi_1 + \pi_2 + \pi_3 \cdots
\end{aligned}$$

$\Rightarrow 0$

$\pi_0 + \pi_1 + \pi_2 + \pi_3 \cdots = 1$

$\Rightarrow$ indeterminate equation

IV. Series and Others

Exercise 12 (Infinite geometric series)

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

$$S = a + ar + ar^2 + ar^3 \ldots$$

$$- \quad r \cdot S = \quad + ar + ar^2 + ar^3 \ldots$$

$$(1-r)S = a.$$

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

Exercise 13 (Finite geometric series)

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

$$
\begin{aligned}
S &= a + ar + ar^2 + ar^3 + \cdots ar^{n-1} \\
- \quad rS &= \quad\; ar + ar^2 + ar^3 + \cdots ar^{n-1} + ar^n
\end{aligned}
$$

$$
\begin{aligned}
(1-r)S &= a - ar^n \\
&= a(1-r^n)
\end{aligned}
$$

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

Exercise 14 (Power series)

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

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

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

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

$$\vdots = \vdots$$

$+$

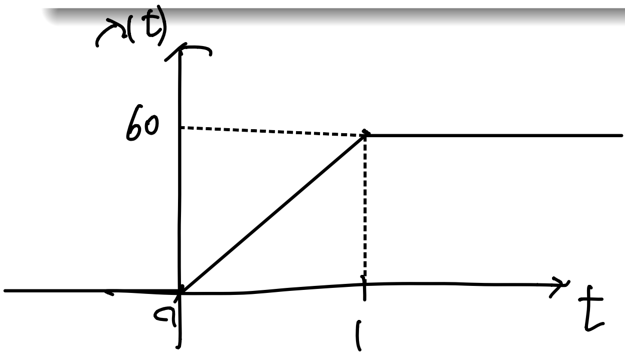
$$r + 2r^2 + 3r^3 + 4r^4 + \cdots = \frac{1}{1-r}\left(r + r^2 + r^3 + \cdots\right)$$

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

Formulation of time varying function

Exercise 15

During the first hour $(0 \leq t \leq 1)$, $\lambda(t)$ *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.*



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

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