機械学習入門

その3

福岡工業大学短期大学部
情報メディア学科
東　昭太朗

アジェンダ

- 機械学習に必要な数学の知識　～基礎編～（続）
  - 微分
  - 指数
  - 対数
- 機械学習に必要な数学の知識　～応用編～
  - シグモイド関数

機械学習で知っておくべき数学　～基礎編～

何を使うのか？

微分（・積分）	線形代数	指数・対数
• 常微分 • 偏微分	• ベクトル • 行列	• 指数関数 • 対数関数 • ネイピア数

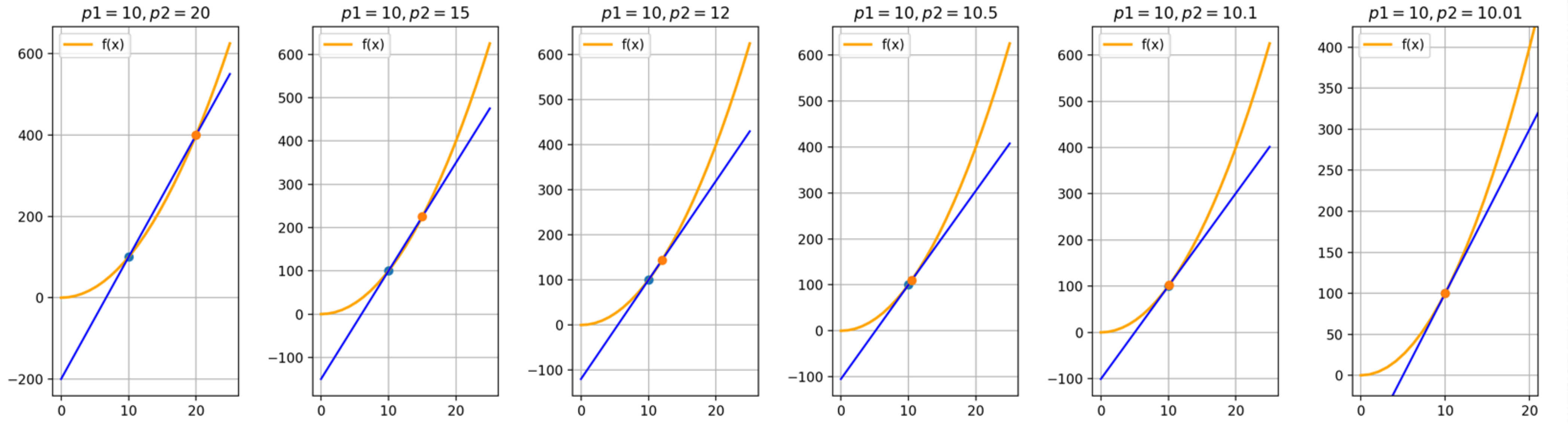
微分とは？

- 変化前と変化後の2点をとり、その変化の割合(平均変化率)を求めることに着想を得ている
- その2点間を**限りなく近づけた**時の値を求めることを**微分する**という
- 任意の点xにおける微分係数を求める関数が導関数となる

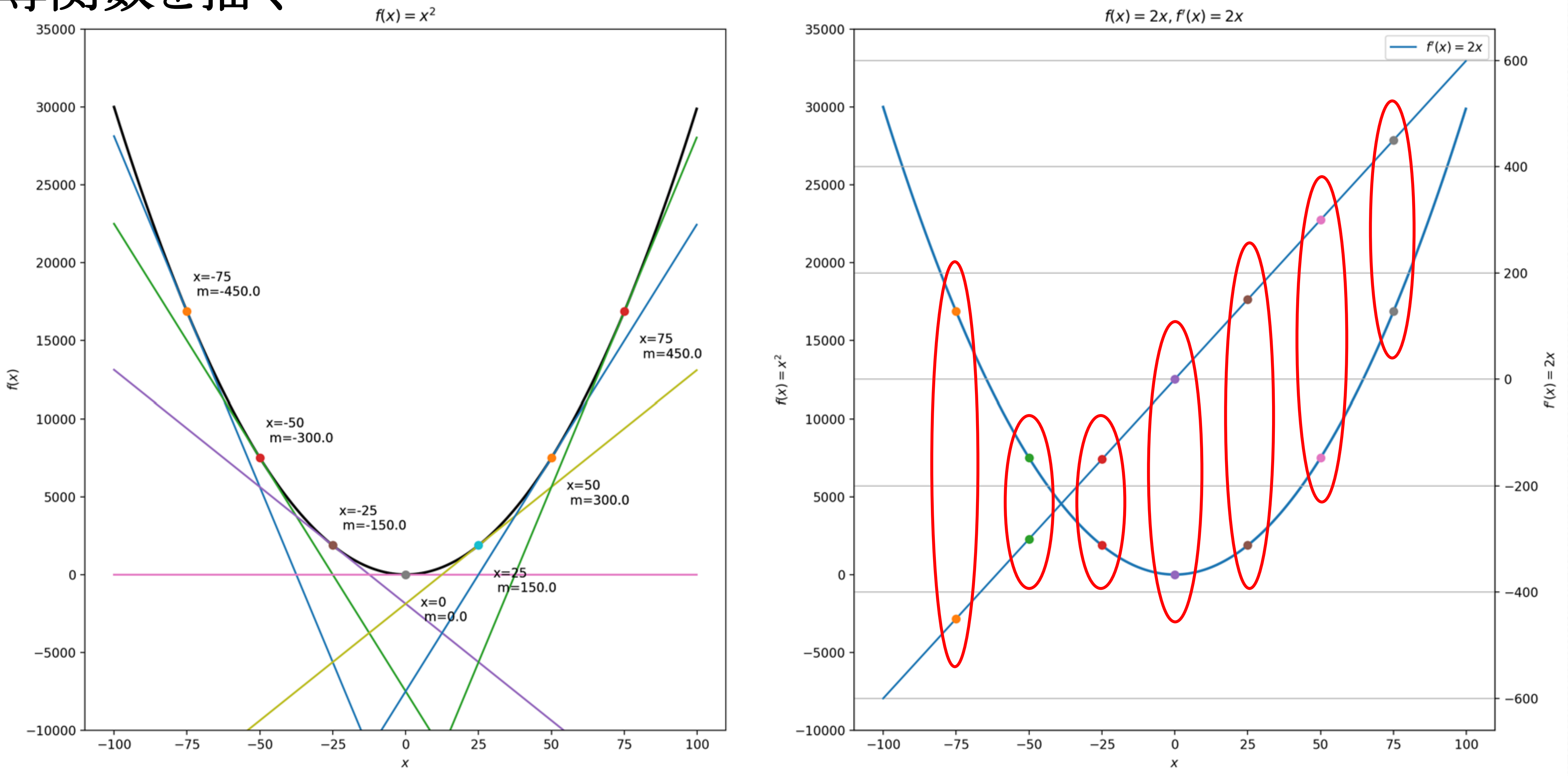
$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \qquad \left(\frac{y\text{の変化量}}{x\text{の変化量}}\right)$$

2点間の距離(差)を限りなく0に近づけていく　＝　極限

$$\lim_{h \to 0}$$



導関数を描く



$f(x) = x^2$
x=-75
m=-450.0
x=-50
m=-300.0
x=-25
m=-150.0
x=0
m=0.0
x=25
m=150.0
x=50
m=300.0
x=75
m=450.0
$f(x)$
x
$f(x) = 2x, f'(x) = 2x$
$f'(x) = 2x$
$f(x) = x^2$
$f'(x) = 2x$
x

ネイピア数

- 超越数と呼ばれる数学の世界の特殊な定数
- 無理数なので無限小数となる
- e と表すのが定石
- 異なる経緯で発見された2つの定義が存在する

$$\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n \qquad \lim_{h \to 0} \frac{e^h - 1}{h} = 1$$

ネイピア数の定義　その1

- きっかけは複利計算
- 発見者はヤコブ・ベルヌーイとされる
- 基本は1年複利（一年ごとに利息がつく）で計算するが、これを1カ月複利、1日複利と短く区切って考えるとどうなるか検討する過程で発見された

単利法 $a(1+nr)=b$

複利法 $a(1+r)^n=b$

a : 初期値（元本）　**n** : 期間（年数）　**r** : 利率（年利）　**b** : 将来価値

元金:100円　年利:12%　の例

$$a(1+r)^n=b$$

1年後　$100(1+0.12)^1=100\times1.12=112.0$

2年後　$100(1+0.12)^2=100\times1.2544=125.4$

3年後　$100(1+0.12)^3=100\times1.404928=140.4$

1カ月複利に直すと・・・

$$100\left(1+\frac{0.12}{12}\right)^1=100\times1.0100=101.00$$

$$100\left(1+\frac{0.12}{12}\right)^2=100\times1.0201=102.01$$

$$100\left(1+\frac{0.12}{12}\right)^3=100\times1.0303=103.03$$

$$\cdots$$

$$100\left(1+\frac{0.12}{12}\right)^{12}=100\times1.1268=112.68$$

単位期間が短い方が得をする？？

更に1日複利にすると・・・

1月	$100\left(1+\frac{0.12}{12}\right)^1 = 100 \times 1.0100 = 101.00$	1日	$100\left(1+\frac{0.12}{365}\right)^1 = 100 \times 1.0003 = 100.03$
2月	$100\left(1+\frac{0.12}{12}\right)^2 = 100 \times 1.0201 = 102.01$	30日	$100\left(1+\frac{0.12}{365}\right)^2 = 100 \times 1.0099 = 100.99$
3月	$100\left(1+\frac{0.12}{12}\right)^3 = 100 \times 1.0303 = 103.03$	90日	$100\left(1+\frac{0.12}{365}\right)^3 = 100 \times 1.0300 = 103.00$
	$\dots$		$\dots$
1年	$100\left(1+\frac{0.12}{12}\right)^1 2 = 100 \times 1.1268 = 112.68$	1年	$100\left(1+\frac{0.12}{365}\right)^3 65 = 100 \times 1.1274 = 112.\underline{74}$

$$\lim_{n\to\infty}\left(1+\frac{r}{n}\right)^{nx}$$

ただし、r:年利率, n:単位期間（n=1⇨1年）, x:年数とする

ここでの議論は単位期間を短くとれば取るほど得をするのかということ

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n}$$

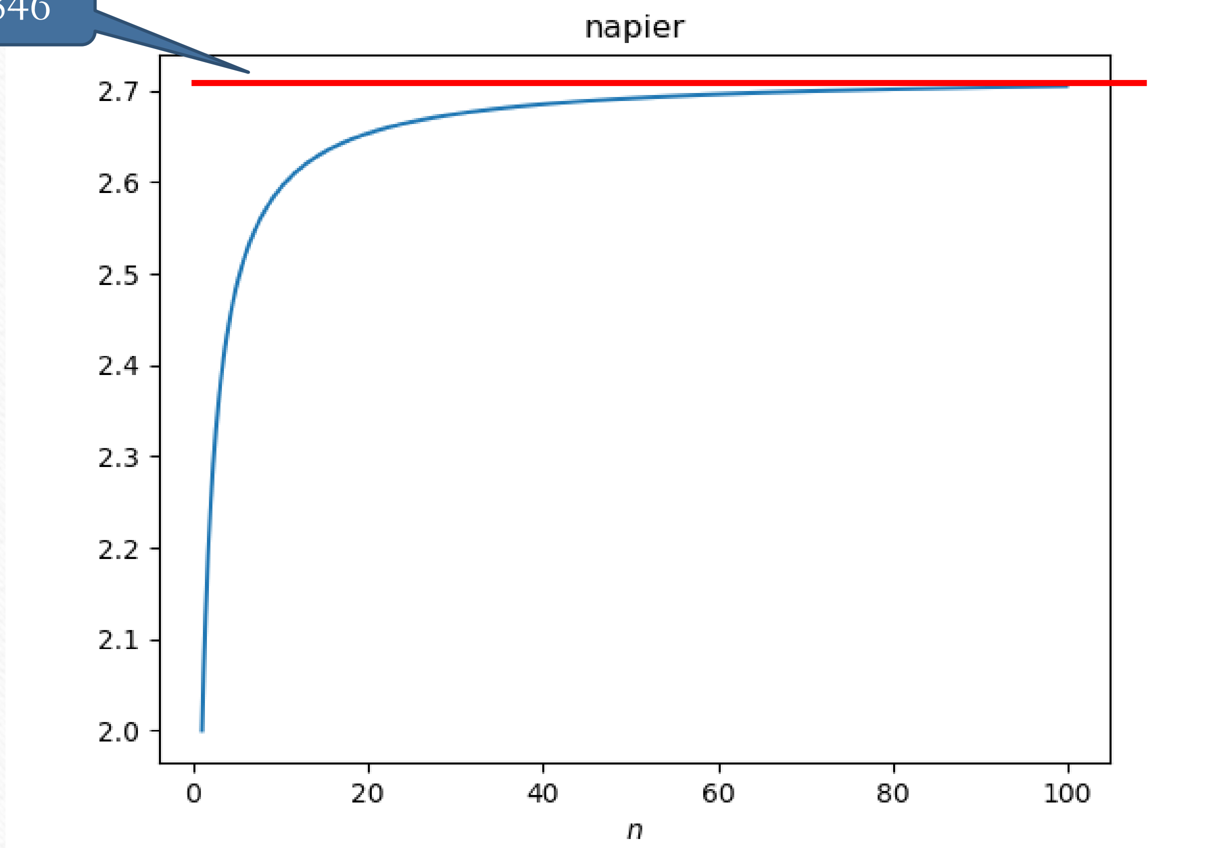
そのほかの変数は排除

2.71828182846

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n}=e$$

但し、n ≧ 1 の実数とする

細かく区切っても無限に良くなるわけではない



1 ÷ 365 = 0.0273973…

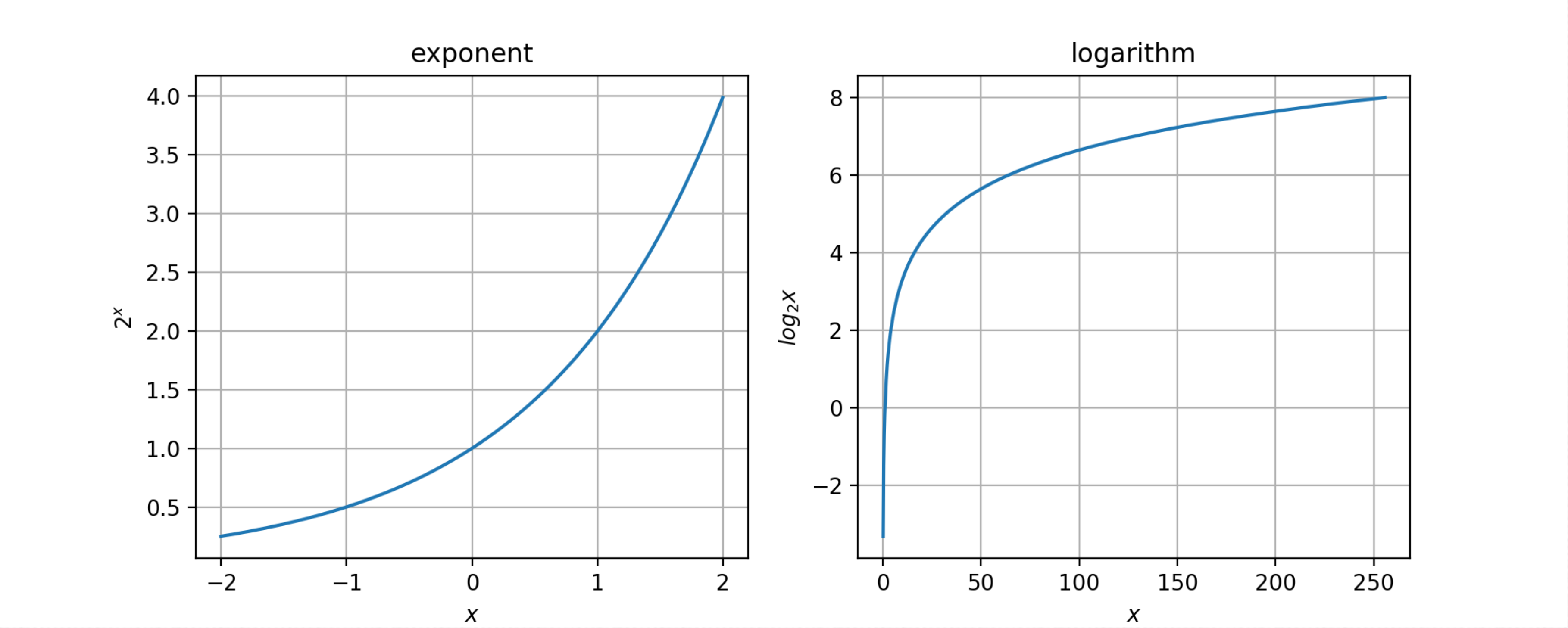
ネイピア数の定義　その2

- 発見者はオイラー
- 対数関数の導関数を導く過程で見出した
- この式を満たす **e** が定義

$$\lim_{h \to 0} \frac{e^h - 1}{h} = 1$$

指数と対数

$$f(x) = a^x \qquad f(x) = \log_a x$$



対数の定義

指数　　対数

$$a^r = P \iff \log_a P = r$$

a を r 回かけると<u>何</u>になるか？

$$\underbrace{a \times a \times \cdots \times a}_{r\,\text{個}} = ?$$

a を<u>何回</u>かけると P になるか？
（**a** を P にする指数は何か？）

$$\underbrace{a \times a \times \cdots \times a}_{?\,\text{個}} = P$$

逆関数の関係

対数の基本性質

$$\log_a M + \log_a N \Leftrightarrow \log_a MN$$

$$\log_a M^p \Leftrightarrow p \log_a M$$

$$a^{\log_a b} = b$$

対数の基本性質 その1

$$\boxed{\log_a M + \log_a N \Leftrightarrow \log_a MN}$$

$a^x = M \cdots (1),\ a^y = N \cdots (2)$ とすると

対数の定義より

$a^x = M,\ a^y = N$

とかける

ここで指数法則より

$a^{x+y} = a^x a^y$

$\therefore a^{x+y} = MN$

この対数をとって

$\log_a MN = x + y$

ここに(1),(2)を代入して

$\log_a MN = \log_a M + \log_a N$

対数の基本性質　その2

$$\boxed{\log_a M^p \Leftrightarrow p\log_a M}$$

$\log_a M = x \cdots (1)$とすると

対数の定義より

$a^x = M$

とかける

ここで両辺をp乗して

$a^{xp} = M^p \cdots (2)$

(2)の対数をとって

$\log_a M^p = xp$

ここに(1)を代入して

$\log_a M^p = p\log_a M$

対数関数の微分

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h} \qquad f(x)=\log_a x$$

$$\frac{d}{dx}\log_a x = \lim_{h\to 0}\frac{\log_a(x+h)-\log_a x}{h}$$

$$= \lim_{h\to 0}\frac{\log_a(x+h)+(-1)\log_a x}{h}$$

基本性質その2

$$= \lim_{h\to 0}\frac{\log_a(x+h)+\log_a x^{-1}}{h}$$

$$= \lim_{h\to 0}\frac{\log_a(x+h)+\log_a\left(\frac{1}{x}\right)}{h}$$

基本性質その1

$$= \lim_{h\to 0}\frac{\log_a\left(\frac{x+h}{x}\right)}{h}$$

$$= \lim_{h\to 0}\frac{1}{h}\log_a\left(\frac{x+h}{x}\right)$$

基本性質その2

$$= \lim_{h\to 0}\log_a\left(\frac{x+h}{x}\right)^{\frac{1}{h}}$$

$$= \lim_{h\to 0}\log_a\left(1+\frac{h}{x}\right)^{\frac{1}{h}}$$

$$= \lim_{h\to 0}\log_a\left(1+\frac{h}{x}\right)^{\frac{1}{x}\frac{x}{h}}$$

基本性質その2

$$= \lim_{h\to 0}\frac{1}{x}\log_a\left(1+\frac{h}{x}\right)^{\frac{x}{h}}$$

$$= \lim_{h\to 0}\frac{1}{x}\log_a\left(1+\frac{1}{\frac{x}{h}}\right)^{\frac{x}{h}}$$

$$= \frac{1}{x}\log_a\lim_{h\to 0}\left(1+\frac{1}{\frac{x}{h}}\right)^{\frac{x}{h}}$$

$$= \frac{1}{x}\log_a e$$

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e$$

指数関数の微分

$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$ $f(x) = a^x$

$$
\begin{aligned}
\frac{d}{dx} a^x &= \lim_{h \to 0} \frac{a^{x+h} - a^x}{h} \\
&= \lim_{h \to 0} \frac{a^x a^h - a^x}{h} \\
&= \lim_{h \to 0} \frac{a^x (a^h - 1)}{h} \\
&= a^x \lim_{h \to 0} \frac{a^h - 1}{h} \\
&= a^x \lim_{h \to 0} \boxed{\frac{a^{\log a^h} - 1}{\log a^h} \frac{\log a^h}{h}}
\end{aligned}
$$

$$\lim_{t \to 0} \frac{e^t - 1}{t} = 1$$

?

$$a^x \log a$$

機械学習に必要な数学の知識　～応用編～

よく使われる関数

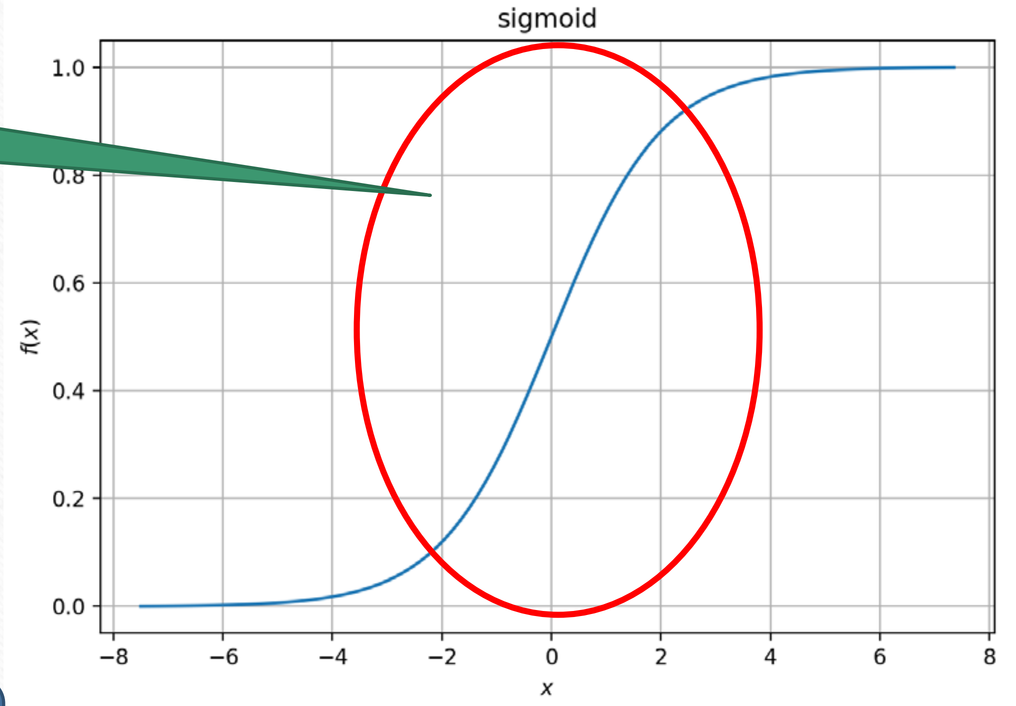
- シグモイド関数
- ソフトマックス関数
- ガウス関数
- Relu関数

シグモイド関数

- あらゆる実数を0から1の範囲に変換する
- その性質上確率を表す
- ニューラルネットワークにおいて、ニューロンの発火現象の特性を表現するための関数として用いられる

連続したグラフ → 微分可能である

$$f(x) = \frac{1}{1+\exp(-x)} = \frac{1}{1+e^{-x}}$$
$$= \frac{1}{1+\frac{1}{e^x}} = \frac{1}{\frac{e^x+1}{e^x}} = \frac{e^x}{e^x+1}$$



分母は分子に1足しただけ！？

シグモイド関数を微分する

リファレンス

- http://www.minemura.org/juken/taisu_seishitsu.html
- https://mathtrain.jp/logseisitsu