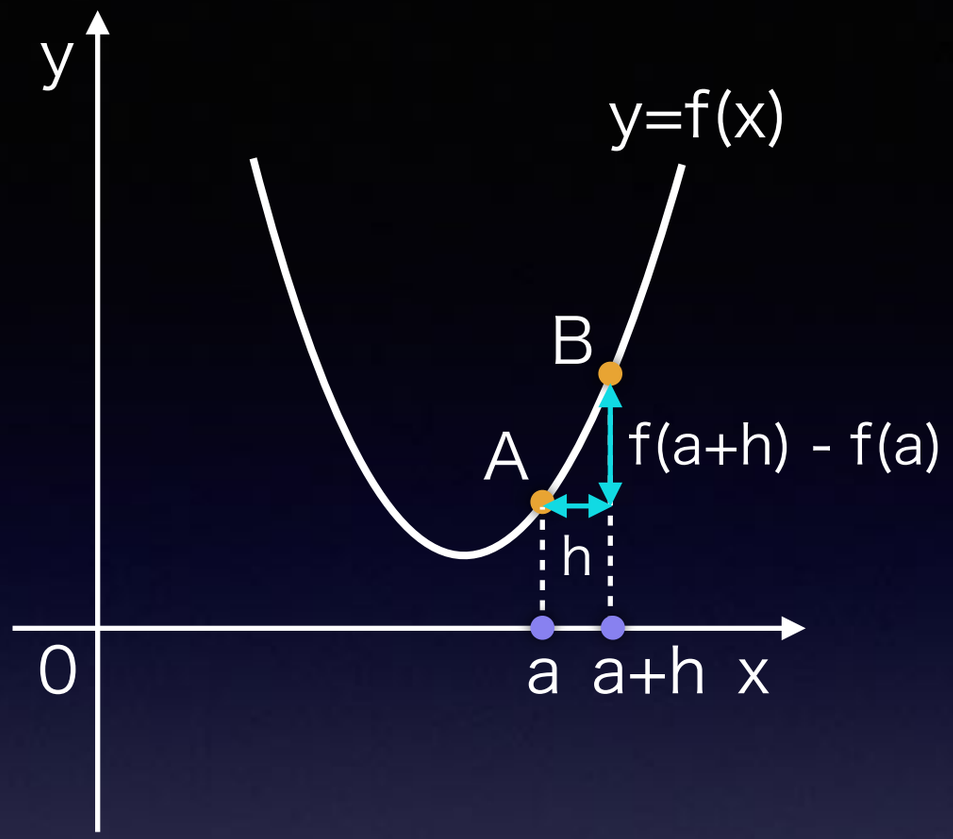
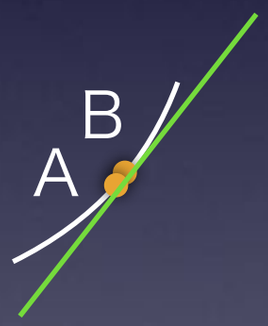
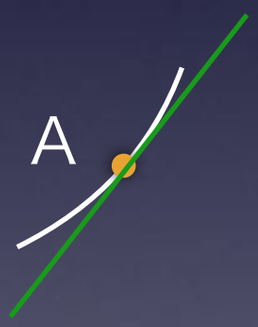




A,Bを通る直線



Aの接線

（Aのみを通る直線）

■ 微分概要

左記のような２次関数y＝f(x)を例に説明する。

この関数上で、x=aの点をA、そこから距離h離れたx=a+hの点をBとする。
すると、2点A,Bを通る直線の傾きは、
(f(a+h) - f(a)) / h
となる。
そして、hを際限なく小さくしていき、０に近づけていく。すると、BはAに一致はしないものの、一致しているとみなしても差し支えないほど近づく。
そうなると、2点A,Bを通る直線は、Aのみを通る直線に一致しているとみなしても差し支えないほど近づく。
y=f(x)に対してAのみ通る直線とは、つまり、y=f(x)にAで接する直線、すなわち接線ということになる。
よって、2点A,Bを通る直線の傾きを接線の傾きとみなすことができる。
以上のように、関数上の2点の内、1点をもう1点に限りなく近づけることで、近づかれる方の点における接線の傾きを求めるのが微分である。

数式上で、微分であることを表すには、
y’=, f’(x) =, dy/dx= などと書く。

■ 偏微分概要

偏微分は、例えば、
z＝f(x,y)・・・変数が2つ 、
z = f(w,x,y)・・・変数が3つ
などの2つ以上の変数を持つ関数において、
その内の1つの変数だけを変数として扱い、
それ以外の変数を定数として扱って微分を行う。
例えば、
$z=x^2+y^2$
という関数でxについて偏微分する場合は、
xは変数扱い、yは定数扱いとなるので、
zは変数がxだけの関数とみなされる。
よって、結果は、
$\partial z / \partial x = 2x$
となる。
yについて偏微分する場合は、x,yの扱いが
逆になり、
$\partial z / \partial y = 2y$
となる。
$\partial z / \partial x$, $\partial z / \partial y$ は、それぞれ、xとyについての
偏微分という意味である。

図形的に理解しようとすると、
$z=x^2+y^2$
のような３次元空間に単純な曲面を持つ関数が適している。
xについての偏微分は、
まず、曲面を任意のyの値のところでx軸と平行に切断した
時の断面を想像すれば良い。
すると、あとは、微分の基本的な話と同じである。
つまり、断面（縦軸：z、横軸：x）上の曲線において、
点Aとそこからわずかに距離hだけ離れた点Bを考え、
BをAに限りなく近づけていけば、点Aにおける接線の
傾きを求めることができる。
同様に、yについての偏微分では、任意のxの値のところで
y軸と平行に切断した時の断面を想像すれば良い。

■ 微分/偏微分の機械学習・深層学習での利用について

機械学習では、ある入力を与えた時に、期待する出力にできるだけ近い出力を出す関数を作成することがある。
そのためには、出力誤差＝(期待する出力ー関数の出力)ができる限り小さくなるように関数の出力を調整する必要がある。
関数の形（項数や各項での変数の次数など）が決まっている場合、関数の出力の調整＝式中の係数の調整である。
例えば、関数に含まれる係数が２個の場合を考えると、係数の値をa , b などとおいて変数として扱い、出力誤差をそれらの関数として表す。
２次元以上の関数の場合、関数の最小値は、関数が描く曲線・曲面の谷底（※関数が谷を持つ場合）であり、そこでは各変数の軸方向の接線の傾きが０である。
よって、a, bについて偏微分を行い、それらの偏導関数が０になる時のa, bの値を求めれば、出力誤差を最小にする係数が求まる。

以上の説明では誤差の最小化に利用されている例を挙げたが、機械学習では、評価値、報酬値などと呼ばれる値を最大化したいために上記とは逆に値を最大化するような係数の値（関数形状の山頂地点の値）を求めることもある。

深層学習では、ニューラルネットワーク内の重みの値を更新する際に偏微分が利用されている。
大まかな説明をすると、その時点での出力誤差に対する重みの影響度の大きさに応じて、重みを更新する。
重みをw、出力誤差関数をEとすると、

$$w_{(更新後)} = w_{(更新前)} - \rho \left(\partial E / \partial w\right)$$

となる。（ρは学習率などと呼ばれる係数）
通常、重みの初期値はランダムな値に設定される。
そこから入力をニューラルネットワークに入れて出力誤差を求め、上記のように重みを更新することを何度も繰り返し、より良い出力を出せるニューラルネットワークに変えていく。

■ 合成関数の微分

$y = (x^2+3x+1)^4$
$y' = 4(x^2+3x+1)^3 * (x^2+3x+1)'$
$= 4(x^2+3x+1)^3 * (2x+3)$
$= 4(2x+3)(x^2+3x+1)^3$

$y = \log(\sin(x^3-2))$
$y' = 1/(\sin(x^3-2)) * (\sin(x^3-2))'$
$= 1/(\sin(x^3-2)) * (\cos(x^3-2)) * (x^3-2)'$
$= 1/\tan(x^3-2) * 2x^2$
$= 2x^2/\tan(x^3-2)$

■ 合成関数の偏微分

$f(x,y) = (x^2+y^2)\sin xy$
$\partial f/\partial x = 2x * \sin xy * (x^2+y^2)\cos xy * y$
$= 2xy(x^2+y^2) * \sin xy * \cos xy$
$= 2xy(x^2+y^2) * (\sin(xy+xy) + \sin(xy-xy))/2$
$= 2xy(x^2+y^2) * \sin(2xy)/2$
$= xy(x^2+y^2)\sin(2xy)$