Lecture 3
Calculus

明玉瑞 Yurui Ming

yrming@gmail.com

声明
Disclaimer

- 本讲义在准备过程中由于时间所限，所用材料来源并未规范标示引用来源。所引材料仅用于教学所用，作者无意侵犯原著者之知识版权，所引材料之知识版权均归原著者所有；若原著者介意之，请联系作者更正及删除。

  The time limit during the preparation of these slides incurs the situation that not all the sources of the used materials (texts or images) are properly referenced or clearly manifested. However, all materials in these slides are solely for teaching and the author is with no intention to infringe the copyright bestowed on the original authors or manufacturers. All credits go to corresponding IP holders. Please address the author for any concern for remedy including deletion.

数列极限
Sequence Limits

- 如果无限序列a_1，a_2，$\cdots$，a_n，$\cdots$，当序数n无限增大时，项a_n无限趋近于固定数A，则我们称序列的极限为A，表示为$\lim\limits_{n\to\infty} a_n = A$。

  If the term a_n in an infinite sequence $a_1, a_2, \cdots, a_n, \cdots$, approaches a fixed number A as the ordinal number n increases without bound, we say that the sequence has limit A. We denote this as $\lim\limits_{n\to\infty} a_n = A$.

- 对于$k \in \mathbb{R}$，$\lim\limits_{n\to\infty} a_n = A$，$\lim\limits_{n\to\infty} b_n = B$，关于极限的下面性质成立：

  The following property of limits holds for $k \in \mathbb{R}$, $\lim\limits_{n\to\infty} a_n = A$, $\lim\limits_{n\to\infty} b_n = B$:

$$\lim_{n\to\infty} ka_n = kA \qquad \lim_{n\to\infty} a_n b_n = AB$$

$$\lim_{n\to\infty}(a_n + b_n) = A + B \qquad \lim_{n\to\infty} \frac{a_n}{b_n} = \frac{A}{B}, (B \neq 0)$$

区间
Intervals

- 变量的取值范围通常是一个区间。符号$[a,b]$表示满足不等式$a \leq x \leq b$的所有实数x的集合，称为闭区间。符号(a,b)表示满足不等式$a < x < b$的所有实数x的集合，称为开区间。符号$[a,b)$表示满足不等式$a \leq x < b$的所有实数x的集合，称为左闭右开区间。符号$(a,b]$，$[a,+\infty)$，$(-\infty,b]$，$(a,+\infty)$，$(-\infty,b)$和$(-\infty,+\infty)$可以用类似的方式定义。

  The region of a variable is usually in an interval. The notation $[a,b]$ denotes all real numbers x that satisfy the inequality $a \leq x \leq b$, and it is called a closed interval. The notation (a,b) denotes all real numbers x that satisfy the inequality $a < x < b$, and it is called an open interval. The notation $[a,b)$ denotes all real numbers x that satisfy the inequality $a \leq x < b$. Notations $(a,b]$, $[a,+\infty)$, $(-\infty,b]$, $(a,+\infty)$, $(-\infty,b)$, and $(-\infty,+\infty)$ can be defined in similar ways.

函数
Functions

- 函数可视为两个变量之间的对应关系。根据这种对应关系，一个变量y的值可以由另一个变量x的值确定，也就是说，y是x的函数。这种对应关系通常表示为$y = f(x)$，字母f表示对应关系。变量x称为自变量，变量y称为因变量。自变量的变化范围称为函数的定义域，因变量的变化范围称为函数的值域，函数$f(x)$在$x = a$处的值表示为$f(a)$。

  A function is a correspondence between two variables. From this correspondence, the value of one variable y can be determined by the value of another variable x. That is, y is a function of x. This correspondence is often denoted as $y = f(x)$, and the letter f represents the correspondence. The variable x is called the independent variable, and the variable y is called the dependent variable. The region of change of independent variables is called the domain of the function. The region of change of dependent variables is called the range of the function. The value of a function $f(x)$ at $x = a$ is denoted as $f(a)$.

函数极限

Function Limits

- 对于函数$y=f(x)$，如果当自变量x趋近点a时$f(x)$的值趋近固定值A，我们说当x趋于a时，函数的极限为A。我们将其表示为$\lim_{x\to a} f(x)=A$。

  For a function $y=f(x)$, if the value of $f(x)$ approaches a fixed number A when the independent variable x approaches a point a, we say that the function has a limit A as x approaches a. We denote this as $\lim_{x\to a} f(x)=A$.

- 若$k\in\mathbb{R}$，函数$f(x)$，$g(x)$，$h(x)$极限存在且有限，则下面性质成立：

  If $k\in\mathbb{R}$, the limits of functions $f(x)$, $g(x)$ and $h(x)$ exist and are finite, the following property holds:

$$\lim_{x\to a} kf(x)=k\lim_{x\to a} f(x)$$

$$\lim_{x\to a}\big(f(x)\pm g(x)\big)=\lim_{x\to a} f(x)\pm\lim_{x\to a} g(x)$$

$$\lim_{x\to a}\big(f(x)\cdot g(x)\big)=\lim_{x\to a} f(x)\cdot\lim_{x\to a} g(x)$$

$$\lim_{x\to a}\frac{f(x)}{g(x)}=\frac{\lim_{n\to\infty} f(x)}{\lim_{x\to a} g(x)},\left(\lim_{x\to a} g(x)\neq 0\right)$$

连续函数
Continuous Functions

- 许多函数是“连续的”。例如，粒子移动的位置x是时间t的连续函数，$x = f(t)$。如果粒子从一个位置移动到另一个位置，它将经过连接两个位置的路径上的每一点。

  Many function are “continuous”. For example, the position x of a moving particle is a continuous function of time t, $x = f(t)$. If a particle moves from one position to another position, it passes every point on a path connecting the two positions.

- 函数$f(t)$在点$t = t_0$处是连续的，如果方程$\lim_{t \to t_0} f(t) = f(t_0)$成立。否则$f(t)$在$t_0$处不连续。如果$x = f(t)$在闭区间$a \leq x \leq b$（或开区间$a < x < b$）的每个点都是连续的，那么我们说$f(t)$是$[a, b]$（或$(a, b)$）上的连续函数。

  A function $f(t)$ is continuous at a point $t = t_0$ if the equation $\lim_{t \to t_0} f(t) = f(t_0)$ holds. Otherwise $f(t)$ is discontinuous at t_0. If $x = f(t)$ is continuous at every point of a closed interval $x = f(t)$ (or an open interval $a < x < b$), then we say that $f(t)$ is a continuous function on $[a, b]$ (or on (a, b)).

夹逼定理
Sandwich Theorem

- 对于序列$\{a_n\}$，$\{b_n\}$，$\{c_n\}$，若对如果$\forall n \in \mathbb{N}$，有$a_n \le c_n \le b_n$且$\lim_{n\to\infty} a_n = A$，$\lim_{n\to\infty} b_n = B$，$\lim_{n\to\infty} c_n = C$，则$A \le C \le B$。

  For sequences $\{a_n\}$, $\{b_n\}$, $\{c_n\}$, if for $\forall n \in \mathbb{N}$, it satisfies that $a_n \le c_n \le b_n$and $\lim_{n\to\infty} a_n = A$, $\lim_{n\to\infty} b_n = B$, $\lim_{n\to\infty} c_n = C$, then $A \le C \le B$.

- 若函数$f(x)$，$g(x)$，$h(x)$极限存在且有限，对任意的x，有$f(x) \le h(x) \le g(x)$，且$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = A$，则有$\lim_{x\to a} h(x) = A$。

  If the limits of functions $f(x)$, $g(x)$ and $h(x)$ exist and are finite, and for arbitrary x, it holds $f(x) \le h(x) \le g(x)$ and $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = A$, then we have $\lim_{x\to a} h(x) = A$.

导数与微分
Derivatives and Differentials

- 让我们回顾一个高中物理学过的例子：假设粒子沿直线移动，距离s和时间t之间的函数关系是$s = s(t)$。则粒子在时间t_0和时间t之间行进的距离是$s(t) - s(t_0)$，比率$\bar{v}(t) = \frac{s(t)-s(t_0)}{t-t_0}$是粒子在时间间隔$[t_0, t]$内的平均速度，是关于自变量$t$的函数。

  Let's review one example in physics in high-school. A particle moves along a straight line. Suppose the relation between the distance s and the time t is a function $s = s(t)$. The distance that the particle traveled between time t_0 and time t is $s(t) - s(t_0)$. The ratio $\bar{v}(t) = \frac{s(t)-s(t_0)}{t-t_0}$ is the mean velocity of the particle in the time interval $[t_0, t]$. It is a function of the independent variable t.

- 由于粒子行进过程中速度可能会时刻改变，所以$\bar{v}(t)$不能表示在时间t_0处的速度。但随着时间间隔$[t_0, t]$变短，$\bar{v}(t)$的值将更接近t_0处的速度，其极限是粒子在t_0处的瞬时速度。

  The speed of the particle during travelling may be different. So $\bar{v}(t)$ can not represent the velocity of the partial at the time t_0. Obviously, as the time interval $[t_0, t]$ gets shorter, the value of $\bar{v}(t)$ will get closer to the velocity at t_0. The limit is the instantaneous velocity of the particle at t_0.

导数与微分
Derivatives and Differentials

- 假设$y=f(x)$是定义在区间$[a,b]$上的函数，$x_0\in[a,b]$。如果极限$\lim\limits_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}$存在，那么我们说函数$f(x)$在$x_0$处是可微的，这个极限称为$f(x)$在$x_0$处的导数，记为$f'(x_0)$或$\left.\frac{dy}{dx}\right|_{x=x_0}$。

  Suppose $y=f(x)$ is a function defined on the closed interval $[a,b]$, and $x_0\in[a,b]$. If the limit $\lim\limits_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}$ exists, then we say that the function $f(x)$ is differentiable at x_0, and this limit is called the derivative of $f(x)$ at x_0, denoted as $f'(x_0)$ or $\left.\frac{dy}{dx}\right|_{x=x_0}$.

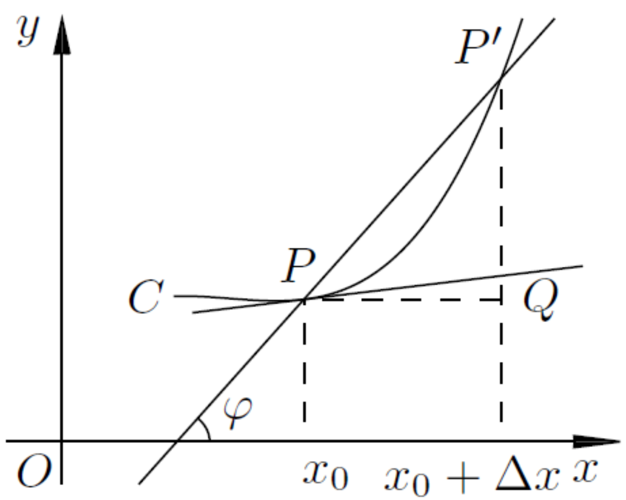
- 显然，$f'(x_0)$的值与点x_0有关，当x_0在a和b之间变化时，$f'(x_0)$会发生变化。如果函数$f(x)$在$[a,b]$的每个点都是可微的，则$f'(x)$是x在同一区间上的一个新的函数，称为$f(x)$的导数 。

  Obviously, the value of $f'(x_0)$ is related to the point x_0. It changes when x_0 changes between a and b. Therefore, if the function $f(x)$ is differentiable at every point of $[a,b]$, then $f'(x)$ is a new function of x on the same interval, called the derivative of $f(x)$.

导数与微分
Derivatives and Differentials

- 下面我们来说一下函数导数的几何意义。如图所示，令曲线C上点P的坐标为$(x_0, f(x_0))$，点P'的坐标为$(x_0+\Delta x, f(x_0+\Delta x))$。连接$P$和$P'$的线段是一条割线，其斜率为$\tan\varphi = \frac{P'Q}{PQ} = \frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}$。当点$P'$沿曲线$C$向点$P$移动时，割线$PP'$绕$P$转，$\Delta x$接近零。因此，割线的斜率接近极限$\lim_{\Delta x\to 0}\tan\varphi = \lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}$。很明显，该极限位置处的线显然是点$P$处曲线$C$的切线。

- We now illustrate the geometric meaning of the derivatives. Suppose the coordinates of a point P on a curve C is $(x_0, f(x_0))$, and the coordinates of another point P' on the same curve is $(x_0+\Delta x, f(x_0+\Delta x))$. The line segment which connects P and P' is a secant line, and its slope is $\tan\varphi = \frac{P'Q}{PQ} = \frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}$. As the point P' moves towards the point P along the curve C, the secant line PP' turns around P, and Δx approaches zero. It follows that, the slope of the secant line approaches the limit $\lim_{\Delta x\to 0}\tan\varphi = \lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}$. Obviously, the line at this limit position is the tangent line to the curve C at point P.



导数与微分
Derivatives and Differentials

- 函数的导数是函数的因变量相对于自变量的变化率。同时，可导或可微是函数的“局部”属性，因为它仅取决于需要求的函数的导数在给定点附近的性质。

  The derivative of a function is the rate of change of the dependent variable of the function with respect to the independent variable. Differentiability is a “local” property of a function since it only depends on the nature of the function nearby the point at which the derivative of the function needs to be found.

- 如果函数$y=f(x)$在点$x=x_0$处可微分，则它在该点是连续的。事实上，在等式$f(x_0+\Delta x)-f(x_0)=\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\cdot\Delta x$的两边都取极限，则$\lim\limits_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\cdot\Delta x=f'(x_0)\cdot\lim\limits_{\Delta x\to 0}\Delta x=0$，即$\lim\limits_{\Delta x\to 0}f(x_0+\Delta x)-f(x_0)=0$，此即$f(x)$在$x_0$连续。

  If a function $y=f(x)$ is differentiable at a point $x=x_0$, then it is continuous at the point. In fact, taking the limit on both sides of the equation $f(x_0+\Delta x)-f(x_0)=\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\cdot\Delta x$, we have $\lim\limits_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}\cdot\Delta x=f'(x_0)\cdot\lim\limits_{\Delta x\to 0}\Delta x=0$, which means $\lim\limits_{\Delta x\to 0}f(x_0+\Delta x)-f(x_0)=0$. This indicates $f(x)$ is continuous at x_0 according to the definition.

导数与微分
Derivatives and Differentials

- 例：求$y=x^n$（n为正整数）点$x=x_0$的导数。

Example: find the derivative of $y=x^n$ (n is a positive integer) at the point $x=x_0$.

$$\lim_{\Delta x\to 0}\frac{(x_0+\Delta x)^n-{x_0}^n}{\Delta x}$$

$$=\lim_{\Delta x\to 0}\frac{{x_0}^n+C_n^1{x_0}^{n-1}\Delta x+C_n^2{x_0}^{n-2}(\Delta x)^2+\cdots+C_n^{n-1}{x_0}^1(\Delta x)^{n-1}+(\Delta x)^n-{x_0}^n}{\Delta x}$$

$$=\lim_{\Delta x\to 0}\left[C_n^1{x_0}^{n-1}+C_n^2{x_0}^{n-2}\Delta x+\cdots+C_n^{n-1}{x_0}^1(\Delta x)^{n-2}+(\Delta x)^{n-1}\right]$$

$$=C_n^1{x_0}^{n-1}=n{x_0}^{n-1}$$

即

$$\frac{dx^n}{dx}=nx^{n-1}$$

导数与微分

Derivatives and Differentials

- 下面代码展示了TensorFlow对多项式函数求导的支持：

  The following code demonstrates the support of TensorFlow for finding derivatives of polynomial functions:

```
import tensorflow as tf
x = tf.Variable(0.0, dtype=tf.float32, name='x')
@tf.function
def y(t):
    global x
    x.assign(t)
    return tf.math.pow(x, 3)
@tf.function
def dy(t):
    global x
    r = y(t)
    return tape.gradient(r, x)

with tf.GradientTape(persistent=True) as tape:
    y_v = y(2.0); dy_v = dy(2.0)

print(f"y = {y_v}, dy = {dy_v}")
```

导数与微分
Derivatives and Differentials

- 例：求$y=\sin x$点$x=x_0$的导数。

Example: find the derivative of $y=\sin x$ at the point $x=x_0$.

$$\lim_{\Delta x\to 0}\frac{\sin(x_0+\Delta x)-\sin(x_0)}{\Delta x}$$
$$=\lim_{\Delta x\to 0}\frac{2\sin\frac{\Delta x}{2}\cos\frac{2x_0+\Delta x}{2}}{\Delta x}$$
$$=\cos\frac{2x_0+\Delta x}{2}\lim_{\Delta x\to 0}\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}$$
$$\lim_{\Delta x\to 0}\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}\overset{\alpha=\frac{\Delta x}{2}}{\Longleftrightarrow}\lim_{\alpha\to 0}\frac{\sin\alpha}{\alpha}$$

下面，求$\lim\limits_{\alpha\to 0}\frac{\sin\alpha}{\alpha}$的极限。

导数与微分

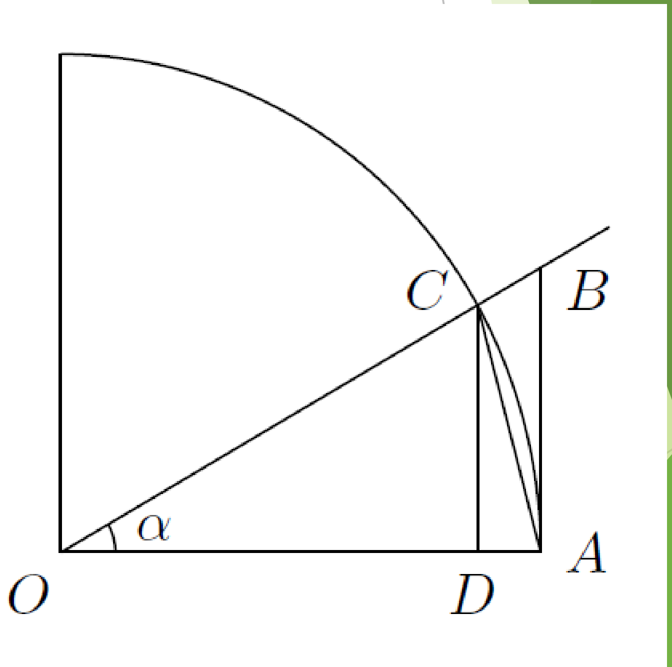
Derivatives and Differentials

- 考虑一个半径为 1 以原点为中心的圆。假设α（$0 \leq \alpha < \frac{\pi}{2}$）是$OA$（$x$轴的正半轴）和直线$OB$之间的夹角。设$C$是圆和$OB$的交点，$CD$是垂直于$OA$的直线。另外，设$BA$是$A$处圆的切线。

  Consider a circle with radius 1 centered at the origin. Suppose α ($0 \leq \alpha < \frac{\pi}{2}$) is an angle between OA (the positive side of x-axis) and the straight line OB. Let C be the intersection of the circle and OB, and CD be a straight line that is perpendicular to OA. Also, let BA be a tangent line of the circle at A.

- 显然，三角形ΔOAB的面积大于扇区OAC的面积，扇形OAC的面积大于三角形ΔODC的面积。

  Obviously, the area of the triangle ΔOAB is greater than the area of the sector OAC, and the area of the sector OAC is greater than the area of the triangle ΔODC.



导数与微分
Derivatives and Differentials

- 三角形ΔOAB的面积为$\frac{1}{2}\tan\alpha$，扇区OAC的面积为$\frac{1}{2}\alpha$，ΔODC的面积为$\frac{1}{2}\sin\alpha$。因此，我们有$\frac{1}{2}\tan\alpha > \frac{1}{2}\alpha > \frac{1}{2}\sin\alpha$，变换得$\frac{1}{\sin\alpha} > \frac{1}{\alpha} > \frac{1}{\tan\alpha}$，即$1 > \frac{\sin\alpha}{\alpha} > \cos\alpha$。

  The area of the triangle ΔOAB is $\frac{1}{2}\tan\alpha$. The area of the sector OAC is $\frac{1}{2}\alpha$. The area of ΔODC is $\frac{1}{2}\sin\alpha$. Thus, we have $\frac{1}{2}\tan\alpha > \frac{1}{2}\alpha > \frac{1}{2}\sin\alpha$. It follows that $\frac{1}{\sin\alpha} > \frac{1}{\alpha} > \frac{1}{\tan\alpha}$ and $1 > \frac{\sin\alpha}{\alpha} > \cos\alpha$.

- 当α趋近零时，不等式的右侧趋近1，因此根据夹逼原理，$\frac{\sin\alpha}{\alpha} \to 1$当$\alpha \to 0$时。

  The right side of this inequality approaches 1 asαapproaches zero, so according the clapping theorem, $\frac{\sin\alpha}{\alpha} \to 1$ as $\alpha \to 0$.

$$\frac{d\sin x}{dx} = \cos x$$

导数与微分
Derivatives and Differentials

- 下面代码展示了TensorFlow对三角函数求导的支持：

  The following code demonstrates the support of TensorFlow for finding derivatives of sinusoidal functions:

```
import numpy as np
import tensorflow as tf
x = tf.Variable(0, dtype=tf.float32, name="x")
@tf.function
def y(t):
    global x
    x.assign(t)
    return tf.math.sin(x)
@tf.function
def dy(t):
    global x
    r = y(t)
    return tape.gradient(r, x)
with tf.GradientTape(persistent=True) as tape:
    y_v_half_pi = y(np.pi / 2)
    dy_v_half_pi = dy(np.pi / 2)
    y_v_quad_pi = y(np.pi / 4)
    dy_v_quad_pi = dy(np.pi / 4)
y = 1.0, dy = -4.371138828673793e-08
y = 0.7071067690849304, dy = 0.7071067690849304
```

导数与微分
Derivatives and Differentials

- 与导数密切相关的概念是微分。假设函数$y = f(x)$在点x处具有导数$f'(x)$，x的增量和导数$f'(x)$的乘积$f'(x)\Delta x$称为函数$y = f(x)$在点x处的微分，表示为$dy = f'(x)\Delta x$ 。

  A concept closely related to derivative is differential. Suppose a function $y = f(x)$ has the derivative $f'(x)$ at a point x. Then the product of an increment of x and the derivative $f'(x)\Delta x$ is called the differential of the function $y = f(x)$ at point x, and denoted as $dy = f'(x)\Delta x$.

- 我们现在解释微分的含义。在一些实际问题中，当函数的自变量值发生变化时，需要找到函数值的变化。例如，当球的半径r增加Δr时，我们可能需要找到球体积的增量。球的体积为$V = \frac{4}{3}\pi r^3$，当r变为$r + \Delta r$时，体积的增量计算如下：

  We now explain the meaning of the differential. In some real problems, the change of a function value needs to be found when the independent variable value of the function changes. For instance, we may need to find the increment of the volume of a ball when its radius r has an increase Δr. The volume of the ball is $V = \frac{4}{3}\pi r^3$. The increment of the volume as r changes to $r + \Delta r$ is calculated as follows:

导数与微分
Derivatives and Differentials

$$\Delta V = \frac{4}{3}\pi(r+\Delta r)^3 - \frac{4}{3}\pi r^3 = \frac{4}{3}\pi(3r^2\Delta r + 3r(\Delta r)^2 + (\Delta r)^3)$$
$$= 4\pi r^2\Delta r + 4\pi r(\Delta r)^2 + \frac{4}{3}\pi(\Delta r)^3$$

- 当Δr的变化很小时，ΔV的增量可由上式右侧的第一项来逼近。我们称V的微分为$4\pi r^2\Delta r$。

  The increment ΔV can be approximated by the first term of the right hand side of the above equation $4\pi r^2\Delta r$, when the change Δr is small. The term $4\pi r^2\Delta r$ is the differential of V.

- 实际上，根据导数的定义，

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x} = \lim_{\Delta x\to 0}\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x} = f'(x_0)$$

  令$\alpha = \frac{\Delta y}{\Delta x} - f'(x_0)$，当$\Delta x \to 0$时，$\alpha \to 0$，即$\lim_{\Delta x\to 0}\alpha = 0$。同时，由$\alpha = \frac{\Delta y}{\Delta x} - f'(x_0)$，变换得$\Delta y = f'(x_0)\Delta x + \alpha\Delta x$。但当$\Delta x$趋近于零时，$\alpha$亦趋近于零，即$\alpha\Delta x$趋近于零的速度比$f'(x_0)\Delta x$要快，因此，当$\Delta x$足够小时，我们记$f'(x_0)\Delta x$为$y$在$x_0$处的微分$dy$，在$\Delta x$足够小时用来逼近$\Delta y$。

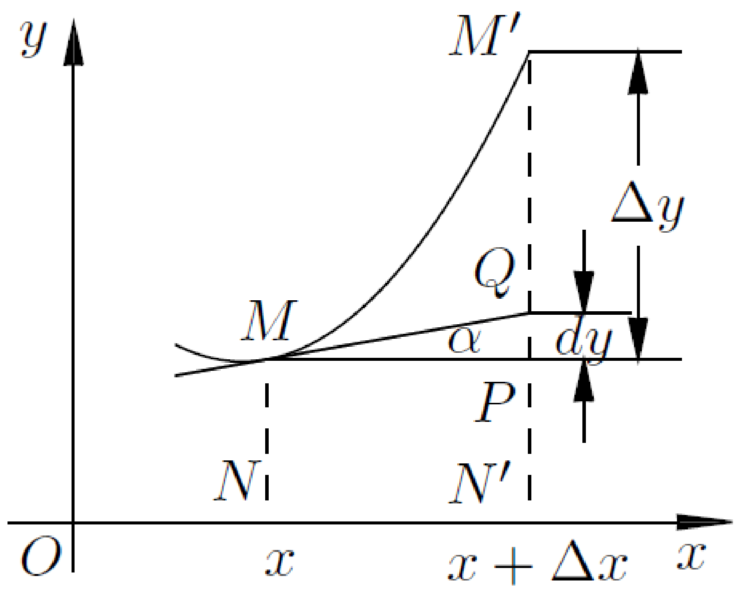
导数与微分
Derivatives and Differentials

- 我们现在研究微分的几何意义。假设$M(x,y)$是曲线$y=f(x)$上的一个点，而$M'(x+\Delta x, y+\Delta y)$是曲线上靠近$M$的另一个点。

  We now study the geometric meaning of differential. Suppose $M(x,y)$ is a point on a curve $y=f(x)$, and $M'(x+\Delta x, y+\Delta y)$ is another point on the curve near M.

- 作垂线MN和$M'N'$垂直于x轴。此外，让M处曲线的切线在Q处与$M'N'$相交.假设α是切线和x轴之间的角度。

  Let vertical lines MN and $M'N'$ be perpendicular to x-axis. Also, let the tangent line to the curve at M intersect $M'N'$ at Q. Suppose α is the angle between the tangent line and the x-axis.



导数与微分
Derivatives and Differentials

- 由上图，我们有：

$$MP = \Delta x$$
$$M'P = M'N' - PN' = f(x+\Delta x) - f(x) = \Delta y$$
$$PQ = MP\tan\alpha = f'(x_0)\Delta x = dy$$

因此，Δy是垂直方向上曲线的变化，dy是垂直方向上切线的变化。

Thus, Δy is the change of the curve on the vertical axis, and dy is the change of the tangent on the vertical axis.

- 考虑函数$y = x$，我们有$f'(x) = x' = 1$。因此，此函数的微分为$dx = 1 \cdot \Delta x = \Delta x$。这意味着，自变量的微分就是变量的变化。因此，我们可以重写$dy = f'(x)\Delta x$为$dy = f'(x)dx$。

Consider the function $y = x$. We have $f'(x) = x' = 1$. Thus, the differential of this function is $dx = 1 \cdot \Delta x = \Delta x$. This implies that, the differential of the independent variable is the change of the variable. Hence we can rewrite $dy = f'(x)\Delta x$ as follows $dy = f'(x)dx$.

导数与微分
Derivatives and Differentials

- 上述的dx和dy分别是自变量和因变量的微分。因此，函数的导数是这两个微分的比值：$f'(x) = dy/dx$。

  Here dx and dy are the differentials of the independent variable and the dependent variable respectively. The derivative of the function is therefore the ratio of these two differentials: $f'(x) = dy/dx$.

- 微分也是函数的“局部”属性。它仅取决于要求的微分的点x附近的函数的性质。

  The differential is also a “local” property of a function. It only depends on the nature of the function near the point x where the differential is to be found.

乘积法则
The Product Rule

- 现在，我们来研究函数乘积的导数。乘积法则也被称为莱布尼茨法则，以戈特弗里德·莱布尼茨的名字命名，他在1684年发现了此规律。这是一个非常重要的法则，因为它允许我们求导更多的函数。令$f(x)=g(x)h(x)$，则：

  We now look at the derivative of a product of functions. The product rule is also called Leibniz rule named after Gottfried Leibniz, who found it in 1684. It is a very important rule because it allows us differentiate many more functions. Let $f(x)=g(x)h(x)$, then we have:

$$
\begin{aligned}
f'(x) &= \lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x} = \lim_{\Delta x\to 0}\frac{g(x+\Delta x)h(x+\Delta x)-g(x)h(x)}{\Delta x}\\
&= \lim_{\Delta x\to 0}\frac{g(x+\Delta x)h(x+\Delta x)-g(x)h(x+\Delta x)+g(x)h(x+\Delta x)-g(x)h(x)}{\Delta x}\\
&= \lim_{\Delta x\to 0}\frac{\big(g(x+\Delta x)-g(x)\big)h(x+\Delta x)+g(x)\big(h(x+\Delta x)-h(x)\big)}{\Delta x}\\
&= \lim_{\Delta x\to 0}\frac{\big(g(x+\Delta x)-g(x)\big)h(x+\Delta x)}{\Delta x}+\lim_{\Delta x\to 0}\frac{g(x)\big(h(x+\Delta x)-h(x)\big)}{\Delta x}\\
&= g'(x)h(x)+g(x)h'(x)
\end{aligned}
$$

乘积法则
The Product Rule

- 由乘积法则，我们可以得到如下商法则。但我们先从一个简单结论推导开始：

  From the product rule, we can easily draw the quotient rule. However, we begin from the deduction of a simple conclusion:

$$\frac{d}{dx}\frac{1}{f(x)} = \lim_{\Delta x \to 0}\frac{\frac{1}{f(x+\Delta x)} - \frac{1}{f(x)}}{\Delta x} = \lim_{\Delta x \to 0}\frac{f(x) - f(x+\Delta x)}{\Delta x f(x+\Delta x) f(x)} = \lim_{\Delta x \to 0}\frac{-f'(x)}{f^2(x)}$$

- 则令$f(x) = g(x)/h(x)$，根据乘积法则与上面结论，我们有：

  Assume $f(x) = g(x)/h(x)$, from the product rule and the conclusion above, we have:

$$f'(x) = \frac{d\left(g(x) \cdot \frac{1}{h(x)}\right)}{dx} = g'(x) \cdot \frac{1}{h(x)} + g(x) \cdot \frac{d}{dx}\frac{1}{h(x)} = g'(x) \cdot \frac{1}{h(x)} + g(x) \cdot \frac{-h'(x)}{h^2(x)}$$
$$= \frac{g'(x)h(x) - g(x)h'(x)}{h^2(x)}$$

链式法则
The Chain Rule

- 如果我们想取复合函数的导数，如$f(x) = \sin(x^3)$，乘积法则不起作用，因为这些函数不是相乘的，而是“链式”的。在该例中，因为我们首先计算x^3，然后对它应用sin。实际上，我们先取在x^3处函数sin的导数，然后将其乘以x^3的导数，答案是$\cos(x^3)\,3x^2$。

  If we want to take the derivative of a composition of functions like $f(x) = \sin(x^3)$, the product rule does not work. The functions are not multiplied but are “chained” in the sense that we evaluate first x^3 then apply sin to it. In order to differentiate, we take the derivative of the function sin evaluated at x^3, then multiply the derivative of the x^3. The answer is $\cos(x^3)\,3x^2$.

- 实际上，令$f(x) = g\big(h(x)\big)$，我们有：

$$f'(x) = \lim_{\Delta x \to 0} \frac{g\big(h(x+\Delta x)\big) - g\big(h(x)\big)}{\Delta x} = \lim_{\Delta x \to 0} \frac{g\big(h(x+\Delta x)\big) - g\big(h(x)\big)}{h(x+\Delta x) - h(x)} \cdot \frac{h(x+\Delta x) - h(x)}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{g\Big(h(x) + \big(h(x+\Delta x) - h(x)\big)\Big) - g\big(h(x)\big)}{h(x+\Delta x) - h(x)} \cdot \frac{h(x+\Delta x) - h(x)}{\Delta x}$$

链式法则
The Chain Rule

- 进一步地，

$$\lim_{\Delta x\to 0}\frac{g\left(h(x)+\big(h(x+\Delta x)-h(x)\big)\right)-g\big(h(x)\big)}{h(x+\Delta x)-h(x)} \overset{\Delta h=h(x+\Delta x)-h(x)}{\Longleftrightarrow} \lim_{\Delta h\to 0}\frac{g(h(x)+\Delta h)-g\big(h(x)\big)}{\Delta h}$$
$$= g'\big(h(x)\big)$$

$$\lim_{\Delta x\to 0}\frac{h(x+\Delta x)-h(x)}{\Delta x}=h'(x)$$

- 综上，我们有：

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

导数与微分
Derivatives and Differentials

- 下面代码展示了TensorFlow对乘积法则与链式法则的支持：

  The following code demonstrates the support of TensorFlow for the product rule and chain rule:

```
import numpy as np
import tensorflow as tf
x = tf.Variable(0, dtype=tf.float32, name="x")
@tf.function
def f(t):
    global x
    x.assign(t)
    return tf.math.sin(x) * tf.math.cos(x)
@tf.function
def f_prime(t):
    global x
    r = f(t)
    return tape.gradient(r, x)
@tf.function
def g(t):
    global x
    x.assign(t)
    return tf.math.sin(tf.math.cos(x))
@tf.function
def g_prime(t):
    global x
    r = g(t)
    return tape.gradient(r, x)
with tf.GradientTape(persistent=True) as tape:
    f_v = f(np.pi / 4); df_v = f_prime(np.pi / 4)
    g_v = g(np.pi / 4); dg_v = g_prime(np.pi / 4)
print(f"f = {f_v}, df = {df_v}")
print(f"g = {g_v}, dg = {dg_v}")
```